ì

1...

Release 3.1A John F. Collins, Biocomputing Research Copyright (c) 1993-1998 University of Edinburgh, U Distribution rights by Oxford Molecular Ltd Research nburgh, U . K.

MPsrch_pp protein - protein database search, using Sat May 13 08:21:01 2000; MasPar time 273.59 Seconds 133.045 Million cell updates/sec Smith-Waterman algorithm

Tabular output not generated

Perfect Description: Score: 3808 MVISKSPFIVLIFSLLLSFA......FLNHKQNTNVIKFTVKASAY

of.

<u>ں</u>

Sequence:

Scoring table: PAM 150 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Listing Match 0% summaries

sptremb112

١,

1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 49.125; Variance 94.306; scale 0.521

and is derived NO. derived is the number of results predicted by chance to have a ster than or equal to the score of the result being printed rived by analysis of the total score distribution.

Result No. 99 110 111 113 114 115 117 119 Score 841 837 895 793 764 724 724 724 726 478 478 478 476 Match 100.0 23.3 22.1 23.1 Length 2346 2346 2346 2346 2346 544 544 544 544 544 544 544 544 544 DB 043358 092W55 092W55 092W55 092W55 022121 041677 041676 041677 041676 040873 040873 040844 040847 04081255 081255 081255 SUMMARIES VICILIN PRECURSOR.
ALPHA SUBUNIT OF BETA
CONVICILIN PRECURSOR.
VICILIN 47KD PROTEIN.
VICILIN-LIKE STORAGE P
VICILIN-LIKE STORAGE P
P54 PROTEIN. GLOBULIN-1
GLOBULIN-1
GLOBULIN-1 GLOBULIN-1 **EMBRYO** GLOBULIN-1 L ALLELE PR ALPHA' SUBUNIT OF BETA BETA SUBUNIT OF BETA C VICILIN PRECURSOR. Description GLOBULIN 1 (FRAGMENT).
1 (FRAGMENT).
1 (FRAGMENT).
1 (FRAGMENT).
1 (FRAGMENT).
1 (FRAGMENT).
1 (FRAGMENT). Pred. No. ; 18e-239 ; 26e-147 2.90e-145 2.90e-146 1.90e-137 1.90e-137 1.92e-139 7.22e-129 7.22e-119 1.90e-114 1.90e-114 1.90e-114 1.90e-114 1.90e-114 1.90e-114 1.90e-69 3.70e-69 3.70e-69 3.70e-69 3.70e-69 3.70e-69 3.70e-69 .00e+00

STIGNING TS 081251 081259 081259 081250 081250 081260 081260 081260 081260 081260 041750 044613 044613 044613 044613 044727 043866 04476 JULY 23.001APC VICILIN (FRAGMENT). 7S STORAGE PROTEIN ALP GLOBULIN (FRAGMENT). VICILIN-LIKE PROTEIN (BETA-CONGLYCININ-ALPHA GLOBULIN (FRAGMENT). GLOBULIN-LIKE PROTEIN GLUTAMINE REPEAT PROTE PHASEOLIN PRECURSOR (F PHASEOLIN PRECURSOR. PHASEOLIN PRECURSOR. VICILIN, 75 GLOBULIN. PHASEOLIN PRECURSOR. GLOBULIN-1 (FRAGMENT). GLOBULIN-1 (FRAGMENT). GLOBULIN-1 (FRAGMENT). GLOBULIN-LIKE PROTEIN GLOBULIN-1 (FRAGMENT). GLOBULIN-1 (FRAGMENT). GLOBULIN-1 S ALLELE PR VICILIN. GLOBULIN-1 (FRAGMENT). GLOBULIN1. GLOBULIN-LIKE PROTEIN ALPHA-PHASEOLIN PRECUR 2.5.92e-69 2.42e-67 1.58e-67 1.09e-66 4.20e-66 4.20e-66 2.30e-66 2.40e-60 2.40e-60 2.50e-60 3.10e-50 3.10e-30 3.10e-30

дb	Db Oy	ОУ	M B Q	SO SO	X D	2 2 2	R R	RT RA	₹ ??	중	8	8 8	80	ဝူ	, i	Ŋ	P A	RESULT ID Q	
121 HENYHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTRFRDEEGNFKILQRFAENSPPLKGIN 180	61 - GĞÖRÜZEREYKEQORQQEEELQRQYOQCOGRCQEQQQGQREQQQCQRKCWEQYKEQERGE 120 	1 MVISKSPETVLIFSLLLSFALLCSGVSAYGRKQYERDPRQQYEQGQRRCESEATEEREQE 60	Query Match 100.0%; Score 3808; DB 10; Length 525; Best Local Similarity 100.0%; Pred. No. 0.00e+00; Matches 525; Conservative 0; Mismatches 0; Lidels 0; Gaps 0;	SIGNAL 1 24 POTENTIAL. CHAIN 25 525 VICILIN. SEQUENCE 525 AA; 60798 MW; 4890EACE CRC32;	Seedstore_7s;	EMBL; X62625; CAA44493.1;	questions about vicilin evolut		TISSUE-LEAVES; MEDLINE; 92288309.	SEQUENCE FROM N.A.	core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Theobroma.	euphyllophytes; Spermatophyta; Magnollophyta; eudicotyledons;		CSV.	U1-NOV-1999 (Trembire1: 12, Last annotation update) VICILIN PRECURSOR.	(TrEMBLrel. 01,	Q43358; 01-NOV-1996 (TrEMBLrel. 01, Created)	ILT 1 Q43358 PRELIMINARY; PRT; 525 AA	

Ş B á DЬ ş В S 밁 Best á ۲ δ g á В ٠ کې Query Q Matches B Ş 문 ç 9 515 210 Q9ZWI3; Q9ZWI3; 01-MAY-1999 01-MAY-1999 PV100. 456 STRAIN-KUROKAWA AMAKURI NANKIN; MEDLINE; 99107919. 150 397 Cucurbita maxima (Pumpkin) (Winter squash),
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyt.
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; 344 eWi 289 AMADA K., SHIMADA T., KONDO M., NISHIMURA M., HARA-NISHIMURA I.; multiple functional protestins are produced by cleaving Asm-Gln bout a single precursor by vaccolar processing enzyme."; Biol. Chem. 274:2563-2570(1999).

Biol. Chem. 274:263-2570(1999).

FOREMER 180(1995): BAA34056:1; -.. 90 31 Local EQUENCE FROM N.A. 481 48] 421 421 Match 361 361 301 301 241 GKYELFFPAGNNKPESYYGAFSYEVLETVFNTQREKLEEILEEQRGQKRQQGQQGMFRKA GEFKDYLSAGGE-SQAYYSVFSNDVLEAALNIPRDKLERIFK-QR----121 HENYHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTRFRDEEGNFKILQRFAENSPPLKGIN R-FQSRYRSDEGHWRVLERFSERSELLKGIKNQRLALLEARPHTFIVPHHLDAECVLLVV 455 RREOERRRRER-EQ-ERR-EREHRGGRDVEDENQR-DPDWRREOERRRREEEEQRE-REWE RSFQTRFRDEEGNFKILQRFAENSPPLKGINDYRLAMFEANPNTFILPHHCDAEAIYFVT GRCQEQQQGQREQQQCQRKCWEQYKEQERGEHENYHNHKKNRSEEEEGQQRNNPYYFPKR -R-EHGRRG-REEORSREDERRRH-EROHGGRSRV-NQVAIRRTEOE--QSNNPYYFOEO NGKGTITFVTHEN---RKOYERDPROQYEOGORKGESHATEENEQECCEORGEREYK-EOOROOEEELOROYOOCO NONGREFEACPEDESQEONMDVAVSAEKLNOGAIEVPHYNSKATEVVEVTDGYGYAQMAC TVYVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNKPESYYGAFSYEVLETVFNTQRE TVYVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNKPESYYGAFSYEVLETVFNTQRE h 33.5%; Scot Similarity 43.6%; Prec 218; Conservative 133; 910 AA; (TrEMBLrel. 10, (TrEMBLrel. 10, (TrEMBLrel. 10, PRELIMINARY; 97315 MW; Score 1275; Pred. No. 5 Last sequence up Created) B9DE560C CRC32; d. No. 5.18e-239; Mismatches 117; PRT; TISSUE-COTYLEDON 018 ₽ update) update) Length Indels RERG - - GKIVRA Tracheophyta; 810; 32; bonds 514 209 325 567 265 149 396 89 343 480 480 420 420 360 360 300 300 240 180 25;

RESULT
AC QOO
AC Ş Ş ₽ 밁 Ş В õ В Matches Query Match Best Local : S ₽ ô 밁 ð 밁 Ş 8 SIGNAL 216 SEQUENCE MENDEL; 11234; Zeama;1188;11234. PFAM; PF00546; Seedstore_7s; 1. 260 141 CHAIN Seed storage protein. EMBL; X59083; CAA41809.1; 01-NOV-1996 01-NOV-1996 Genetics 129:863-872(1991)
-!- MISCELLANEOUS: THE THRI
THE DESIGNATION L, I, ÷ MEDLINE: 92090707.
BELANGER F.C., KRIZ A.L.;
"Molecular basis for allelic gene. STRAIN-W64A 97 87 SEQUENCE FROM N.A. Zea mays (Maize).
Eukaryota; Viridiplantae;
euphyllophytes; Spermatop) 39 GLOBULIN-1 L Poaceae; 003865; 003865 502 747 442 883 LPVNSPGKYELFFPAGNNKPESYYGAFSYEVLETVFNTQREKLEETLEEQRGQKRQQGQQ HTISVPGEFQFFFGPGGRNPESFLSSFSKSIQRAAYKTSSDRLERLFG-RHGQDKGIIVR PROTEINS, RESPECTIVELY.

MISCELLANEOUS: THREE PROTEIN-PROCESSING S
FORMATION OF THE MATURE PROTEIN FROM THE 385 628 326 568 CDAEAIYFYTNGKGTITFYTHENKESYNYQRGTYYSVPAGSTYYYYSQDNQEKLTIAVLA SIMILARITY: TO OTHER 7S CONVICILIN, CONGLYCININ, TDAHCICYVAEGEGVVTTIENGERRSYTIKQGHVFVAPAGAVTYLANTDGRKKLVITKIL NN-PYYFPKRRSFQTRFRDEEGNFKILQRFAENSPPLKGINDYRLAMFEANPNTFILPHH QCVRRCEDRPWHQR--PRCLEQCREEEKEKRQERSRHEADDRSGEGSSEDEREQEKEKQK QGQGRCQEQQQGQREQQQCQRK@WEQYKE--QERGEHENYHNHKKNRSEEE-E-G-Q-QR PRODUCT 161; GRPFFLNH-KQNTNVIKFTV GRENIMNELDREAKELAFNV 766 EEREERS-GRFERVAGRLSEGGVLVIPAGHPIAIMASPNENLRLVGFGINAENNHRNFLA EESEEETFGEFQQVKAPLSPGDVFVAPAGHAVTFFASKDQPLNAVAFGLNAQNNQRIFLA SVVDIKOGGMAVPHFNSRATWVVFVSEGAGSFEMACPHIOSSOWQRGRREEERHWRREEE 687 SQEQLRALSQRATSVRKGSRGVRAPIKLESQTPVYNNQYGQMFEACPDEFPQLRRTDVAT Similarity 35.2%; 161; Conservative KPEQIRAISGOATSPRING-GERLAINLLSOSPYYSNONGREFEACPEDESOFONMDVAV 20 582 AA; (Tremblrel. 01, Created)
(Tremblrel. 01, Last sequence update)
(Tremblrel. 12, Last annotation updat
L ALLELE PRECURSOR (GLB1-L) (7S-LIKE). PRELIMINARY; Spermatophyta; 22.3%; 35.2%; 582 66162 19 I, AND S FOR MW. 118; Mismatches 149; Streptophyta; Embryophyta; Trach hyta; Magnoliophyta; Liliopsida; SEED STO.). Score 848; DB 10; Pred. No. 9.26e-147; GLOBULIN-1 7DEFC426 POTENTIAL. polymorphism of the S ST COMMONLY S PRT; STORAGE 582 CRC32; PROTEINS ₹ OCCURING GLB1 ALLELES INTERMEDIATE, AND SMAI STEPS OCCUR IN THE PRIMARY TRANSLATION update) Length 582; Indels maize (PHASEOLIN, VICILIN, Tracheophyta; sida; Poales; Globulin-1 29; SMALL Gaps 140 259 199 96 HAVE 50 746 24;

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 11:35:26 2000; MasPar time 66.80 Seconds 101.831 Million cell updates/sec

Run on:

Tabular output not generated

Description: Perfect Score: Sequence: >US-09-331-631-7

Title:

1 MVISKSPFIVLIFSLLLSFA.....FLNH

FLNHKQNTNVIKFTVKASAY 525 - Scaldnot

Scoring table: PAM 150 Gap 11

131253 seqs, 12956647 residues

Searched:

Database:

Post-processing: Minimum Match 0% Listing first 45 summaries

a-issued 1:5A_COMB 2:5B_COMB 3:PCT_COMB 4:backfiles1

Statistics: Mean 33.485; Variance 165.024; scale 0.203

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

110 110 110 110 1110 1110 1110 1110 11	Result
3654 18254 18058 18058 7766 7776 7776 173 173 1147 147 147 120 120 120 121 121 122 123 123 123 123 123 123 123	Score
2248.00 244.00 244.00 244.00 244.00 244.00 244.00 244.00 240.00 240.00 240.00 240.00 240.00 240.00 240.00 240.00	Query Match
566 566 587 571 4210 4211 1162 1162 11898 1898 1898 1898 1893 1303 303 303 303 303 303 303 303 303 3	Length
11111111111111111111111111111111111111	DB.
US-07-955- US-07-955- US-07-955- US-07-955- US-07-955- US-08-728-	ID
Sequence	Description
2, Applicatio 22, Applicati 23, Applicati 24, Applicati 25, Applicati 26, Applicati 27, Applicati 27, Applicati 34, Applicati 44, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 48, Applicati 49, Applicati 49, Applicati 40, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 48, Applicati 49, Applicati 40, Applicati 40, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 48, Applicati 49, Applicati 40, Applicati 40, Applicati 41, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 48, Applicati 49, Applicati 40, Applicati 40, Applicati 41, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 47, Applicati 48, Applicati 49, Applicati 49, Applicati 40, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 47, Applicati 48, Applicati 49, Applicati 49, Applicati 40, Applicati 40, Applicati 41, Applicati 41, Applicati 42, Applicati 43, Applicati 44, Applicati 45, Applicati 46, Applicati 47, Applicati 47, Applicati 48, Applicati 49, Applicati 40, Applicati 40, Applicati 41, Applicati 41	ion
0.00e+00 1.70e-144 2.12e-54 3.12e-54 3.12e-54 3.00e-25 1.62e-04 7.46e-04 7.46e-04 7.90e-01 9.90e-01 9.90e-01 9.90e-01 9.90e-01 9.90e-01 9.12e-01 1.27e-02 4.73e-02 9.90e-01 9.90e-01 9.16e-00 1.16e-00 1.16e-00 2.16e+00 4.60e+00	Pred. No.

5 4	4	Δ	42	41	40	ω 9	38 8	37	36	35	34	ယ	32	31	30	29	28	27	26	25	24	
96	96	96	96	95	95	95	95	95	97	94	94	. 94	100	104	104	103	104	106	108	108	110	
																	2.7					
1876	1876	1135	905	559	542	542	542	542	361	162	162	78	614	737	737	360	303	903	678	162	2703	
21	N	N	2	۲,	سو		w	_	;	N	2	N	w			N	<u>_</u>	N	w	_	1	
US-08-609-	- 609 - 80 - SU	US-08-574-	US-08-574-	US-08-313-	US-08-328-	US-07-814-	PCT-US92-1	US-08-258-	US-08-415-	US-08-461-	US-08-486-	US-08-465-	PCT-US95-0	US-08-185-	US-08-185-	US-08-531-	US-08-185-	US-08-853-	PCT-US93-0	US-08-244-	US-08-185-	
Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	
12,	28,	7,	9	14,	00	13,	13,			63	5	ű	12	4.	~	Ņ	,	,	٠ (Ĺ	19,	
Applicati	Applicati	Applicatio	Applicatio	Applicati	Applicatio	Applicati	, Applicati	Applicati	Applicatio	Applicati	Applicati	Applicati	Applicati	Applicatio	Applicatio	Applicatio	Applicatio	Applicacio	Applicatio	Applicatio	Applicati	
3.80e+01	3.80e+01	3.80e+01		4.39e+01	4.39e+01	4.39e+01	4.39e+U1	4.39e+01	3.28e+01	5.08e+01	5.08e+01	5.08e+01	Z. 116+01		1.106+01	1.35e+U1	1.16e+U1	3.0/6+00	0.556+00	6 330+00	4.66e+00	

Дb	ργ	DЪ	Qy	υb	M G C	တ္တ ၁၁	გგ	გგ	8	გგ	86	88	88	CC	88	8	88	88	88	?×	DE:	X D	×č	×	RESULT		
121 HENYHNHKKNRSEEEEGOORNNDYYFPKRRSFOTRFRDEEGNFKILORFAENSPPLKGIN 180	61 OCEORCEREYKEOOROOEEELOROYOOCOGRCQEOQOGOREOQOCORKCWEOYKEOERGE 120	61 QCEQRCEREYKEQQRQQEEELQRQYQQCQGRCQEQQQGGREQQQQCQRKCWEQYKEQERGE 120	1 MVISKSPETVLIFSLLLLSFALLCSGVSAYGRKQYERDPRQQYEQCQRRGESEATEEREQE 60	1 MATSKSALTATION TO THE TOTAL THE TOTAL TO T	Query Match 96.0%; Score 3654; DB 1; Length 566; Best Local Similarity 96.7%; Pred. No. 0.00e+00; Matches 504; Conservative 10; Mismatches 6; Indels 1; Gaps 1;	MOLECULE TYPE: protein SEQUENCE 566 AA; 65613 MW; 1529150 CN;	TYPE: amino acid TOPOLOGY: linear	SEQUENCE CHARACTERISTICS: LENGTH: 566 amino acids	INFORMATION FOR SEQ ID NO: 2:	CLASSIFICATION: 435	APPLICATION NUMBER: US/07/955,905A	SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)		MEDIUM TYPE: Floppy disk	NUMBER OF SEQUENCES: 20 COMPUTER READABLE FORM:	INVENTION:	NVENTION:	GENERAL INFORMATION:	5770433	Sequence 2. Application US/07955905A	Sequence 2, Application US/07955905A		XXXXXX		LT 1 US-07-955-905A-2 STANDARD; PRT; 566 AA.	ALIGNMENTS	

6

•

VINCOSOPPINION PROPERTY Best Loc Matches Query Match . . Ş 밁 밁 Š 밁 ş 8 Ş B . 9 SEQUENCE Local Similarity Sequence 22, Application US/07955905A US-07-955-905A-22 4. 12: HENYHNHKKNRSEEEEGQQRNNPYYFPKHRSFQTRFRDEEGNFKILQRFAENSPPLKGIN 180 1 MVISKSPFIVLIFSLLLSFALLCSGVSAYGRKOYERDDRQQYEQCQRRCESEATBEREQE 60 INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: 481 481 421 361 361 GENERAL INFORMATION: FEATURE MOLECULE TYPE: protein ORIGINAL SOURCE: SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) MEDIUM TYPE: Floppy TITLE OF INVENTION: TITLE OF INVENTION: NUMBER OF SEQUENCES: NAME/KEY: Protein LOCATION: 1.566 OTHER INFORMATION: ORGANISM: APPLICATION NUMBER: US/07/955,905A FILING DATE: 21-JAN-1993 CLASSIFICATION: 435 OPLNAVAFGLNAQNNORIFLAGRPFFLNHKQN-TNVIKFTV 520 OPLNAVAFGLNAONNORIFLAGKKNLVROMDSEAKELSFGV 521 PHLSRQSQGSQSGRQDRREQEEESEEETFGEFQQVKAPLSPGDVFVAPAGHAVTFFASKD PHLSRQSQGSQSGRQDRREQEEESEEEFEGEFQQVKAPLSPGDVFVAPAGHAVTFFASKD NQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATFVVFVTDGYGYAQMAC NONGREFEACPEDESOFONMDVAVSAEKLNGGAIEVPHYNSKATEVVEVTDGYGYAOMAC 420 COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS LENGTH: KSPFIVLIFSLLLSFALLCSGVSAYGRKQYERDPRQQYEQCQRRCESEATEEREQE NO. 22, Application o. 5770433 566 AA: 65613 MW: 1529150 CN; Conservative amino acid 566 amino acids Theobroma linear 96.0%; 96.7%; Floppy disk STANDARD; PRECURSOR 28 /note- "67 kD Precursor Protein" RECOMBINANT 47 AND cacao US/07955905A Score 3654; DB 1; Pred. No. 0.00e+00; Mismatches PRT; 31 kD COCOA PROTEINS 566 AA Length 566 Indels ۳. Gaps 60 480 480 420 360 360 300 300 240

```
Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ٤
    SEQUENCE
                                                                                                                                                                                                                                                                                                                 Sequence 23, Application US/07955905A
                                                                                                                                                                                                                                                                                                                                                              XXXXXX
                                                                                                                 CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                 US-07-955-905A-23
                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                                                                                                                                              481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30,1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                            FEATURE:
                                                                MOLECULE TYPE: protein ORIGINAL SOURCE:
                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
                                                                                                                                                                                                                    COMPUTER READABLE FORM:
                                                                                                                                                                                                                                TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                      NAME/KEY:
LOCATION:
OTHER INFORMATION: /note- "Vicilin from ICE 587 AA; 69497 MW; 1625753 CN;
                                                     ORGANISM:
                                                                                  TOPOLOGY: linear
                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                           TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                 QPLNAVAFGLNAQNNQRIFLAGKKNLVROMDSEAKELSFGV 521
                                                                                                        LENGTH:
                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                     QPLNAVAFGLNAQNNQRIFLAGRPFFLNHKQN-TNVIKFTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PHLSRQSQGSQSGRODRREQEEESEEETFGEEQQVKAPLSPGDVFVAPAGHAVTFFASKD 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NONGREFEACPEDESOFONMDVAVSAFKLNOGAIFVPHYNSKATEVVFVTDGYGYAQMAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHLSROSOGSOSGRODRREQEEESEEETFGEFQQVKAPLSPGDVFVAPAGHAVTFFASKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATFVVFVTDGYGYAQMAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KLEEILEEQRGQKRQQGQQGMFRKAKPEQIRAISQQATSPRHRGGERLAINLLSQSPVYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KLEEILEEQRGQKRQQGQQGMFRKAKPEQIRAISQQATSPRHRGGERLAINLLSQSPVYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTITFVTHENKESYNVQRGTVVSVPAGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTITFVTHENKESYNVORGTVVSVPAGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HENYHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTRFRDEEGNFKILQRFAENSPPLKGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QCEORCEREYKEQQHQQEEELQHQYQQCQGRCQEQQGGREQQQCQRKCWEQYKEQERGE 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HENYHNHKKNRSEEEEGGORNNPYYFPKRRSFOTRFRDEEGNFKILORFAENSPPLKGIN
                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                23,
                                                                                                                                                                                                                                                                                       3, Application 5770433
                                                                                                       587 amino acids
                                Protein
                                                    Gossypium hirsutum
                                                                                                                                             UMBER: US/07/955,905A
21-JAN-1993
                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                 PRECURSOR
28
                                                                                                                                                                                                                                                    RECOMBINANT 47
                                                                                                                                                                                                                                                                                              US/07955905#
                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                   AND
                                                                                                                                                                                                                                                  3
                                                                                                                                                                                                                                                                                                                                                                             587
         <u>ဂ</u>
                                                                                                                                                                                                                                                  õ
                                                                                                                                                                           #
                                                                                                                                                                                                                                                                                                                                                                             ۶
                                                                                                                                                                                                                                               COCOA PROTEINS
                                                                                                                                                                          (EPO)
                                                                                                                                                                                                                                                AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     300
```

:

(MT)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm Sat May 13 11:45:14 2000: MasPar time 58.52 Seconds 130.618 Million cell updates/sec

Tabular output not generated

Run on:

Scoring table: Description: Perfect Score: Sequence: Title: Gap 11 PAM 150 >US-09-331-631-(1-590) from US 0509331631.pep

1 MVRNKSACVVLLFSLFLSFG..... of 4)

FNSNPQESYFVSRQRQRASE 590 - SEQ1 & NO:81

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Statistics: Database: a-issued 1:5A_COMB 2:5B_COMB 3:PCT_COMB 4:backfiles1 Mean 34.226; Variance 174.925; scale 0.196

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB.	ID	Description	Pred. No.
	4119	95.1	587	۱ -	US-07-955-	•	
2 '	2101	48.5	566	_	us-07-955-	Sequence 22, Applicat	·
w	2101	48.5	566	–	us-07-955-	Sequence 2, Application	. ~
•	968	22.3	605	_	us-07-955-	Sequence 24, Applicat	٠,
υn .	832	19.2	571	-	US-07-955-	Sequence 25, Applicat	1 2.626-56
σ, ι	776	17.9	410	-	ò	Sequence 26, Applicat	- μ
7	493	11.4	421	—	US-07-955-	Sequence 27, Applicat	
30 ·	223	.5 .1	1162	N	US-08-728-	Sequence 2. Application	۰α
، ص	166	ა	1898	N	-008-800-	Sequence 94, Applicat	
10	166	3. 8	1898	۲	US-08-056-	Sequence 94, Applicat	1 1.256.03
= :	162	3.7	788	2	US-08-918-	Sequence 4, Application	
12	138	J. 2	186	N	US-08-557-	Sequence 43, Applicat	, <u>, </u>
L i	127	2.9	361	٢	US-08-415-	Sequence 4, Application	
14	123	2.8	737	-	US-08-185-	Sequence 4, Applicatio	070.00
15	123	2.8	737	٢	US-08-185-	Sequence 2, Application	_
16	119	2.7	434	-	US-08-337-	>	
17	119	2.7	2237	N	US-08-223-	•	٠,
18	119	2.7	2237	Ь	US-08-455-	•	٠,
19	119	2.7	2339	٢	US-08-455-	•	۰,
20	119	2.7	2339	N	US-08-223-	Sequence 47, Applicat	n)-
21	112	2.6	357	٢	US-08-910-	Sequence 2, Application	١.
22	112	د .	357	-	US-08-552-	Sequence 2, Application	n y
23	112	2.6	357	w	PCT-US95-0	Sequence 2, Application	

Š 밁

Matches Best Local Query Match

ch 95.1%; 1 Similarity 95.1%; 561; Conservative

Score 4119; DB 1; Pred. No. 0.00e+00; 17; Mismatches 9

Length 587; Indels

ω ··

Gaps

<u>ب</u>

45	44	43	42	41	0	39	38	37	36	<u>ა</u>	34	33	32	31	30	23) N	ر م	27	26	25		2	
107	107	109	109	109	107	109	109	109	110	110	109	109	109	112	717	, j.	, r.	112	112	112	717		113	
2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	N.			ر ا د		2.6			y .	2.6	
387	387	369	356	356	303	303	303	303	240	240	156	156	156	44.0	1	21.	3144	914	914	369		360	369	
۳	-	N	_		٠,	۰,	'n	_	w	ب	-			۰ ۱	٠,	٠.	N	<u>م</u> ــو	_	L	,	_	N	
US-08-329-	05-08-539-	-T66-80-SD	016-80-50	05-08-552-	08-08-180	05-08-105-	US-08-460-	US-08-459-	PCT-US95-0	05-08-261-	US-08-552-	-016-80-50	PCI-0395-0	000000000000000000000000000000000000000	00 00 153	115-08-246-	US-08-457-	US-08-484-	US-08-484-	PC1-0291-0		115-07-688-	US-08-474-	
sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	o ducino	Company	Sequence	Sequence	Sequence	Sequence	Sequence		Sequence	Sequence	
,	,	, ,	٠,	, ,	,	ν ,	'n	,	, ,	3 0	, . ,	9 4	,	0 9	٠.	σ.	42	٧,	,	, ,	2	32	32	
Therroatro	2001100110	APPLICATIO	, , , , , , , , , , , , , , , , , , , ,	A DO LOCAL	100110011	2001100110	Applicació	Applicacio	, Applicact	, Applicaci	200110011	2001100110	2001101010	Anni Catio	Applicatio	Applicatio	Applicati	Applicacio	A PATTOCKO	2004400	Anni Cati	Applicati	Applicati	
	1 160+01	1 160+01	8 55e+00		8 650+00		8 650+00	8 550+00	8 650	7 470+00		8 650+00	8 650+00	8.65e+00	5.57e+00	5.57e+00		1			5.57e+00	5.5/e+00	5.5/e+00	

ALIGNMENTS

RESULT ID US Sequence 23, Application US/07955905A XXXXXX US-07-955-905A-23 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/
FILING DATE: 21-JAN-199
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid SEQUENCE Sequence 23, Application US/07955905A Patent No. 5770433 Patent No. GENERAL INFORMATION: FEATURE:
NAME/KEY: Protein
LOCATION: 1.587
LOTHER INFORMATION: /note- "Vicilin from G. hirsutum"
OTHER INFORMATION: /note- "Vicilin from G. hirsutum"
UENCE 587 AA; 69497 MW; 1625753 CN; MOLECULE TYPE: protein original source: TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND TITLE OF INVENTION: PRECURSOR COMPUTER READABLE FORM: NUMBER OF SEQUENCES: APPLICANT: MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, ORGANISM: TOPOLOGY: 587 amino acids Gossypium hirsutum linear 21-JAN-1993 STANDARD; US/07/955,905A 23: PRT; Version #1.25 (EPO) 587 AA

```
á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ś
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 22, Application US/07955905A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         XXXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-07-955-905A-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           '61 CEEŠČKSQYGEKDQQ-RHRPEDPQKRYEECĞĞECRQQEEKQRPQCQQRCIKRFEQQQQQS 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 KIYLVTNGRGTLTFLTHENKESYNIVPGVVVKVPAGSTVYLANQDNKEKLIIAVLHRPVN
                                                                                                                                                                                                                                                                                                                                                                  Sequence 22, Application US/07955905A Patent No. 5770433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             541 VAGKINHVRQWDSQAKELAFGVSSRLVDEIFNSNPQESYFVSRQRQRASE 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         538 VAGKINHVRQWDSQAKELAFGVSSRLVDEIFNNNPQESYFVSRQRQRASE 587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481 EERRSGQYRKIRSRLSRGDIFVVPANFPVTFVASQNQNLRMTGFGLYNQNINPDHNQRIF 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420 SALQLNQGSIFVPHYNSKATFVVLVNEGNGYVEMVSPHLPRQSSFEEEEEQ--QQEQEQE 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360 ASQEQIRALSQEATSPREKSGERFAFNLLYRTPRYSNQNGRFYEACPREFRQLSDINVTV 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                             CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
                                                                                                                SEQUENCE CHARACTERISTICS LENGTH: 566 amino acid
                                                                                                                                                                                                      SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA:
                                                       MOLECULE TYPE: protein ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                         TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NPROFEEFFPAGSORPOSYLRAFSREILEPAFNTRSEQLDELFGGROSHRROOGGGMFRK 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KTYLVTNGRGTLTFLTHENKESYNVVPGVVVRVPÅGSTVYLANQDNKEKLIIAVLHRPVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PFHFHRRSFOSRFREEHGNFRVLORFASRHPILRGINEFRLSILEANPWTEVLPHHCDAE 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QRQFQECQOHCHQQEQRPEKKQQCVRECREKYQENPWRGEREEEAEEEETEEGEQEQSHN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QRQFQECQOHCHQQEQRPERKQQCVAEGRERYQENPWRREREEEAEEEETEEGEQEQSHN 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CEESCKSQYGEKDQQQRHRPEDPQRRYEECQQECRQQEERQQPQCQQRCLKRFEQEQQQS 120
                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NPGQFEEFFPAGSQRPQSYLRAFSREILEPAFNTRSEQLDELFGGRQSRRRQQGQGMFRK 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PFHFHRRSFQSRFREEHGNFRVLQRFASRHPILRGINEFRLSILEANPNTFVLPHHCDAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EERRSGOYRKIRSQLSRGDIFVVPANFPVTFVASQNONLRWTGFGLYNQNINPDHNQRIF 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SALQLNQGSIFVPHYNSKATFVILVTEGNGYAEMVSPHLPRQSSYEEEEEEDEEEEDEEEEDEE 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ASQEQIRALSQEATSPREKSGERFAFULLSQTPRYSNQNGRFFEACPPEERQLRDINVTV 420
                                                                                                                                                                                                                              COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                       TOPOLOGY:
                                                                                                                                                                           APPLICATION NUMBER: US/U
FILING DATE: 21-JAN-1993
                                           ORGANISM:
                                                                                                                                                                                                                                                               MEDIUM TYPE:
 CATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                566 amino acids
                                           Theobroma cacao
                                                                                     linear
Protein
1..566
                                                                                                                                                                                                                                                              Floppy disk
                                                                                                                                                                                                                                                                                                           PRECURSOR
                                                                                                                                                                                                                                                                                                                        RECOMBINANT 47 AND 31 kD COCOA PROTEINS
                                                                                                                                                                                           US/07/955,905A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ζ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      299
```

```
GGGGGGGGGGGGGGGGXEXH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ŷ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 57.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/07955905A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-07-955-905A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                   XXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           552
                                                                                                                                                                                                                                                                                                            Patent No.
                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/07955905t Patent No. 5770433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         511 DSEAKELSFGVPSKLVDNIFN-NPDESYFMSFSQQRQRRDE 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              396 VPHYNSKATFVVFVTDGYGYAQMACPHLSRQSQGSQSGRQDRREQEEESEEETFGEFQQV 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336 QATSPRHRGGERLAINLLSQSPVYSNQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIF 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      276 NNKPESYYGAFSYEVLETVFNTQREKLEEILEEQRGQKRQQGQQGMFRKAKPEQIRAISQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     253 TFLTHENKESYNIVPGVVVKVPAGSTVYLANQDNKEKLIIAVLHRPVNNPGQFEEFFPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    216 TFVTHENKESYNVQRGTVVSVPAGSTVYVVSQDNQEKLTIAVLALPVNSPGKYELFFPAG 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       136 QRPEKKQQCVRECREKYQENPWRGEREE-EAEEEE-TEEGEQEQSHNPFHFH-RRSFQSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37 DPROQYEQCORRCESEATEEREQEQCEQRCEREYKEQQRQQEEELQRQYQQCQGRCQEQQ 96
              INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82 DPQRRYEECQQEC-RQQ-EERQQPQCQQRCLKRF-EQE-QQQS--QRQFQECQQHCHQQE 135
                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                  COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin-Release #1.0, Version
CURRENT APPLICATION DATA:
                                                                                                                                                                                       COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                       TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RSRLSRGDIFVVPANFPVTFVASQNQNLRMTGFGLYNQNINPDHNQRIFVAGKINHVRQW 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KAPLSPGDVFVAPAGHAVTFFASKDQPLNAVAFGLNAQN-----NQRIFLAGKKNLVRQM 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VPHYNSKATFVILVTEGNGYAEMVSPHLPRQSSYEEEEEEEDEEEEQEQEEERRSGQYRKI 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EATSPREKSGERFAFNLLSQTPRYSNQNGRFFEACPPEFRQLRDINVTVSALQLNQGSIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SQRPQSYLRAFSREILEPAFNTRSEQLDELFGGRQSRRRQQGQ-GMFRKASQEQIRALSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FREEHGNERVLORFASRHPILRGINEFRLSILEANPNTFVLPHHCDAEKIYLVTNGRGTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRDEEGNFKILORFAENSPPLKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTI 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QGQREQQQCQRKCWEQYKEQE-RGEHENYHNHKKNRSEEEEEGQQRNNPYYFPKRRSFQTR 155
                                                                                 APPLICATION NUMBER: US/07 FILING DATE: 21-JAN-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OTHEK INFORMATION: /note- *67 kD Precursor Protein*
ICE 566 AA; 65613 MW; 1529150 CN;
 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative 104; Mismatches 97;
                                                                                                                                                                                      Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                           PRECURSOR
                                                                                                                                                                                                                                                            RECOMBINANT 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 2101; DB 1;
Pred. No. 7.38e-162;
                                                                                                      US/07/955,905A
                                                                                                                                                                                                                                                            AND 31 kD COCOA PROTEINS AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 566;
                                                                                                                                        *1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         312
```

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 08:26:05 2000; MasPar time 33.77 Seconds 733.342 Million cell updates/sec

Tabular output not generated >US-09-331-631-7 (1-525) from US09331631.peg

Description: Perfect Score: Sequence: 3808 1 MVISKSPFIVLIFSLLLSFA... (1 of 3)

.FLNHKQNTNVIKFTVKASAY 525

LONGID33

PAM 150 Gap 11

Scoring table:

Searched: 142080 segs, 47172406 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

pir62 1:pir1 2:pir2 3:pir3 4:pir4

Database:

Statistics: Mean 50.054; Variance 101.476; scale 0.493

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	BB	ID	Description	Pred. No.
_ :	3654	96.0	566	2	S22477	vicilin precursor - c	0.00e+00
N	1919	50.4	605	N	90	lobulin type	0.00e+00
ω	1910	0	509	N	S08059	globulin	0.00e+00
۵	1853	48.7	588	_	FWCNAB	5	0
ر.	881	23.1	573	N	A53234	globulin-1S, GLB1S -	3.82e-145
6	848	22.3	582	N	B53234	vicilin-like storage	13
7	830	21.8	433	N	S00567	vicilin precursor (cl	÷
8	827	21.7	439	-	FWSYCB	beta-conglycinin beta	ů
9	826	21.7	459	2	S08505		
10	810	21.3	605		FWSYBA		. <u>.</u> .
11	805	21.1	605	N	S20007		L
12	776	20.4	571	N	S00566	convicilin precursor	12
13	771	20.2	463	N	A27288	vicilin precursor - f	•
14	771	20.2	463	N	S06309	vicilin precursor (cl	9e-
15	764	20.1	438	2	S35757	vicilin, 47K - garden	1.53e-121
16	732	19.2	410	-	EMPMVB	vicilin B precursor -	
17	724		448	-	S24756	vicilin-like storage	1.59e-113
18	702	18.4	524	N	JQ1730	62K sucrose-binding p	5e-1
19	697		483	2	T06459	62K sucrose-binding p	3e-
20	682	17.9		Ν	S15675	globulin-2 precursor	4 C
21	644	16.9		2	S00281	canavalin – sword bea	7 e
22	643	16.9	445	N	JQ2264	canavalin - jack bean	2.16e-97
23	634	16.6	639	2	B24810	beta-conglycinin alph	1.31e-95

		_	_																		
G	4	ັພ	N	41	ö	39	8	37	36	5	4	ພ	ະ	31	ö	8	28	27	8	5	24
171	173	170	170	183	188	191	195	222	230	286	365	391	418	426	434	435	466	470	476	531	593
4 .5	<u>۵</u> ۰.	4 .5	4.5	4.8	4.9	5.0	5.1	5.8	6.0	7.5	9.6	10.3	11.0	11.2	11.4	11.4	12.2	12.3	12.5	13.9	15.6
1094	708	395	395	1038	184	136	613	124	218	240	296	540	430	414	421	275	407	436	236	TYO	637
N	N	Ν	N	N		N	N	۰	2	2	N	N	N	2	N	ب	N	ν	2		ν,
849313	A53185	S43344	S35565	T02634	FSFB	T02870	S27770	FWPMV4	S16334	T06399	T06572	S21825	S10156	S60190	A24810	FWPMVA	T02258	A23498	T01662	197705	S35221
protein kinase - slim	G-box-binding factor	sex-determining prote	sex-determining prote	rep protein homolog -	phaseolin - kidney be	globulin 2 precursor	hypothetical protein	vicilin, 14K componen	beta-conglycinin alph	7S storage protein al	convicilin precursor	vicilin-like storage	alpha-phaseolin precu	vicilin - Zamia iuriu	5	•	globulini - maize	לַ	globulin-1 - maize (1	6	, eg
8.456-10	4.13e-10	1.21e-09	1.21e-09	1.11e-11	1.76e-12	5.81e-13	1.31e-13	4.48e-18	1.98e-19	3.338-29	1.23e-43	1.6/e-48	1.336-33	4.01e-00	1.20e-36	/./5e-5/	8. 98e-05	1.05e-05	1.0/6-04	1010	1.65e-87

Db 121 HENYHN Qy 121 HENYHN	Db 61 QCEQRC Qy 61 QCEQRC	Db 1 MVISKS 	Query Match Best Local Similarity Matches 504; Conse	FEATURE 1-24 25-566 SUMMARY	GENETICS #introns CLASSIFICATION KEYWORDS	##TOSS-re	,	#cross-refere	#authors #journal #title	ACCESSIONS REFERENCE	RESULT 1 ENTRY TITLE ORGANISM DATE
HENYHNHKKNRSEEDEGQQRNNPYYFPKRRSFQTRFRDEEGNFKILQRFAENSPPLKGIN 180 	OCEORCEREYKEOOROOEEELOROYOOCOGRCOEOQOGOREOOOCORKCWEQYKEQERGE 120 	MVISKSPFIVLIFSLLLSFALLCSGVSAYGRKQYERDPRQQYEQCORRCESEATEEREQE 60 	h 96.0%; Score 3654; DB 2; Length 566; Similarity 96.7%; Pred. No. 0.00e+00; 504; Conservative+ 10; Mismatches 6; Indels 1; Gaps 1;	#domain signal sequence #status predicted #label SIG\ #product, vicilin #status predicted #label MAT #length 566 #molecular-weight 65513 #checksum 3693	211/1; 269/3; 296/3; 391/3; 502/1 #superfamily glycinin seed; storage protein	m I	m 1		MCHENTY, L.; FILEZ, F.J. Plant Mol. Biol. (1992) 18:1173-1176 Comparison of the structure and nucleotide sequence of vicilin genes of cocoa and cotton raise questions about	\$22477; \$22478; \$18105; \$22050 \$22477	S22477 *type complete vicilin precursor - cacao *formal_name Theobroma cacao *common_name cacao 30-Sep-1993 *sequence_revision 30-Sep-1993 *text_change

AGGENPOSYLRIFSREILEAVFNTRSEQLDELPGGRQSHRRQQGQ-GMFRKASQEQ	252 TVTFVTHENKESYNVVPGVVVRIPAGSTVYLANQDNREKLTIAVLHRPVNNPGOE -	Y 154 TRENDEEGNEKILORFAENSPPLKGINDYRLAMFEANPHTEILPHHCDAEAIYE	OREQUOCORKCHEQYKEQE - RGEHEN - YHNHKKNRSEEEEGQQRNNPYYFII	PER-KQQCVKECREQYQEDPWKGERENKWREEEEEESDEGEQQQRNNPYYFH	PORRYODCROHCQQEERRLRPHCEQSCREQY-EKQQQQQPDKQFKECQQRQ	Overy Match 50.4%; Score 1919; DB 2: Length 605; Best Local Similarity 57.0%; Pred. No. 0.00e+00; Matches 269; Conservative 101; Mismatches 87; Indels 15	MARY #length 605 #molecular-weight 71049 #checksum	#Acomain signal sequence #status predicted # #product alpha-globulin type A #status pred	ION #Superfamily glycinin	status not compared wi	alpha globulin (vicilin) genes of cottonseed. ession S06398	urnal Plant Mol. Biol. (1987) 9:533-546 Developmental biochemistry of cottonseed embryog	#authors Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure II	SSIONS S06398	*formal_name Gossypium hirsutum *common_name up 31-Mar-1990 *sequence_revision 31-Mar-1990 *tex	S06398 alpha-c E_NAMES seed st	2	TNVIKETV 5	PLNAVAFGLNAQNNQRIFLAGKK	21 PHLSRQSQGSQSGRQDRREQEEESEEETFGEFQQVKAPLSPGDVFVAPAGHAV	421 PHLSROSOGSOSGRODRREOEEEEEEEEFGEFOOVKAPLSPGDVFVAPAGHAV	36	361 NONGREFEACE		Db 301 KLEEILEEQRGQKRQQGQQGMFRKAKPEQIRAISQQATSPRHRGGERLAINLLSQ	DD 241 TYYVYSQDNQEKLTIAVLALPVNS9GKYELFFPAGNNKPESYYGAFSYEVLETVE	TO BE THE BOTTOM OF THE CONTRACT TO THE PROPERTY OF THE PROPER	SANCTORONAS SANCTO	・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	191 DVDI AKRTANDUTTI DUDODATA TATANDA
Q-GMFRKASQE QGMFRKAKPE	/LHRPVNNDGOEQKFFP 311 - :: /LALPVNSPGKYELFFP 273	HCDAE	RNNPYYEPKRRSEQ 15	RNNPYYFH-RRS	5	ength 605; Indels 1	checks	predicted *status pre		anslatio	nomic organization o	seed embryogenesis and	.A.; Dure III		on_name u -1990 *te	land cotto		520	521	FFASKD 48	- - 4	4 6	FVVFVTDGYGYAOMAC 42	GERLAINLLSQSPVYS 36	RLAINLLSQSPVYS 36	YGAFSYEVLETVFNTQRE 300 YGAFSYEVLETVFNTQRE 300	14 0 VOT 4 4 0 4 L VOO - 74		2000	TWICKBACE

```
Š
                                        밁
                                                                    Ş
                                                                                                      8
                                                                                                                                   ş
                                                                                                                                                                       B
                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                        DP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSIONS
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
ENTRY
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ٤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALTERNATE_NAMES ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       #journal
#title
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       #accession
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         #authors
                              411 RSRLSRGDIFVVPANFPVTFVASQNQNLRMTGFGLYNON 449
                                                                              351 VPHYNSKATFVILVTEGNGYAEMVSPHLPRQSSYEEEEEEEDEEEDQEQEEERRSGQYRKI 410
                                                           396 VPHYNSKATEVVEVTDGYGYAQMACPHLSRQSQGSQSGRQDRREQEEESEEETFGEFQQV
                                                                                                                          336 OATSPRHRGGERLAINLLSOSPYYSNONGRFFEACPEDFSQFONMDVAVSAFKLNOGAIF 395
                                                                                                                                               291 EATSPREKSGERFAFNLLSQTPRYSNQNGRFFEACPPEFRQLRDINVTVSALQLNQGSIF 350
                                                                                                                                                                                                                276 NNKPESYYGAFSYEVLETVFNTQREKLEEILEEQRGQKRQQGQQGMFRKAKPEQIRAISQ 335
                                                                                                                                                                                                                                                            216 TFVTHENKESYNVORGTVVSVPAGSTVYVVSQDNQEKLTIAVLALPVNSPGKYELFFPAG
                                                                                                                                                                                                                                                                                                       172
                                                                                                                                                                                                                                                                                                                             156 FRDEEGNFKILQRFAENSPPLKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTI 215
                                                                                                                                                                                                                                                                                                                                                                      112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        **residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   **molecule_type DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    371
                                                                                                                                                                                                                                                                                                                                                                                                        97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         334
                                                                                                                                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 DPQRRYEECQQEC-RQQ-EERQQPQCQQRCLKRF-EQE-QQQS--QRQFQECQQHCHQQE 54
KAPLSPGDVEVAPAGHAVTFFASKDOPLNAVAFGLNAQN 494
                                                                                                                                                                                                                                                                               TFLTHENKESYNIVPGVVVKVPAGSTVYLANODNKEKLIIAVLHRPVNNPGOFEEFFPAG 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                FREEHGNERVLQRFASRHPILRGINEFRLSILEANPNTFVLPHHCDAEKIYLVTNGRGTL 171
                                                                                                                                                                                                                                                                                                                                                                                              QGQREQQQCQRKCWEQYKEQE-RGEHENYHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTR 155
                                                                                                                                                                                                                                                                                                                                                                                                                              QRPEKKQQCVRECKEKYQENPWKGEKEE-EAEEEE-TEEGEQEQSHNPFHFH-RRSFQSR 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SQGATSPKGKGSEGYAFNLLSQTPRYSNQNGRFYEACPRNFQQQLREVDSSVVAFEINKG 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AIFVPHYNSKATEVVEVTDGYGYAQMACPHLSRQSQGSQSGRQDRREQEEESEEETFGEF 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SQQATSPRHRGGERLAINLLSQSPVYSNQNGRFFEACPEDFSQ-FQNMDVAVSAFKLNQG 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        261;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     h 50.2%;
Similarity 56.9%;
261; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     *superfamily glycinin
*length 509 *checksum 9476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L. Plant Mol. Biol. (1987) 9:533-546

Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Sequences and genomic organization of the alpha globulin (vicilin) genes of cottonseed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S08059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    #formal_name Gossypium hirsutum #common_name upland cotton
31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
30-Sep-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seed storage protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  S06398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S08059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1-509 ##label CHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   not compared with conceptual translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 1910; DB 2; I
Pred. No. 0.00e+00;
95; Mismatches 92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 509;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 11; Gaps
                                                                                                                                                                                                                                                              275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
```

Statistics: Database: Post-processing: Minimum Match 0% Listing first 45 summaries Description: Perfect Score: Result Searched: Scoring table: Sequence: Title: Tabular output Run on: MPsrch_pp Ž O Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. Score Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd protein - protein database search, using Smith-Waterman algorithm not generated 100.0 33.1 21.2 21.2 21.2 20.2 Match Length 705-09-331-631-7 (81-140) from US09331631 pep (3 of 3) 471 1 LOROYOQCOGRCQEQQQGORHENYHWHKKNRSEEEEGQQR 60 PAM 150 Gap 11 Mean 23.260; Variance 99.417; a-issued 1:5A_COMB 2:5B_COMB 3:PCT_COMB 4:backfiles1 Sat May 13 14:47:27 2000: 131253 seqs, 12956647 residues 566 566 587 1898 1898 303 303 303 303 1162 788 176 303 344 404 404 401 605 683 737 361 DB US-07-955US-07-955US-07-955US-08-800US-08-056US-08-460US-08-460US-08-459US-08-459US-08-459US-08-459US-08-415US-08-822US-08-823US-08-823US-08-823US-08-823-US-08-531-US-08-453-US-08-453-US-08-942-US-07-955-5210183-3 SUMMARIES Sequence 2, Application Sequence 23, Application Sequence 24, Application Sequence 24, Application Sequence 24, Application Sequence 2, Application Sequence 2, Application Sequence 2, Application Sequence 4, Application Sequence 4, Application Sequence 3, Application Sequence 3, Application Sequence 3, Application Sequence 2, Application Sequence 24, Appli MasPar time 52.66 Seconds 14.761 Million cell updates/sec Description Sequence 4, Sequence 2, scale 0.234 Applicatio 4.67e-36 4.67e-36 3.81e-06 2.79e-01 7.23e-01 7.23e-01 7.23e-01 7.23e-00 2.70e+00 2.70e+00 3.91e+00 5.65e+00 5.65e+00 1.18e+01 (MT)

Applicatio

24 79 16.8 559 1 US-08-13) Sequence 41, Appl. 27 78 16.6 76 4 S27309-10 Patent No. 5273901-226 78 16.6 76 4 S27309-10 Patent No. 5273901-226 78 16.6 76 4 S27309-10 Patent No. 5273901-226 78 16.6 76 2 US-08-23 Sequence 47, Appl. 27 16.3 76 2 US-08-23 Sequence 55, Appl. 27 16.3 76 2 US-08-24 Sequence 55, Appl. 27 16.3 77 2 US-08-264 Sequence 55, Appl. 27 16.3 77 2 US-08-264 Sequence 55, Appl. 27 16.3 77 2 US-08-264 Sequence 55, Appl. 27 16.3 77 2 US-08-265 Sequence 55, Appl. 27 16.3 77 2 US-08-267 Sequence 64, Appl. 27 16.3 98 2 US-08-267 Sequence 64, Appl. 27 16.5 98 20 US-08-267 Sequence 65, Appl. 28 20 US-08-267 Sequence 67 16.5 98 20 US-08-267 Sequence 68 16.5 98 20 US-08-267 Sequence 69 16.5 98 20 US-08-267 S	ID RE	Qy dd		XEXEXEX	RES	
16. 8 559 1 US-08-313- Sequence 14. APPLIANT 16.3 76 4 5273901-11 Patent No. 54877091 18 16.6 76 4 5273901-11 Patent No. 54877091 16.3 76 2 US-08-481- Sequence 47. Applification 16.3 77 2 US-08-481- Sequence 55. Applification 17. 2 US-08-181- Sequence 64. Applification 17. 2 US-08-181- Sequence 4. Applification 17. 2 US-08-181- Sequence 4. Applification 17. 2 US-08-181- Sequence 4. Applification 17. 2 US-08-181- Sequence 63. Applification 18. 2 US-08-181- Sequence 63. Applification 18. 2 US-08-181- Sequence 63. Applification 18.		· cht	EQU IN Seq		ULT US-	44444000000000000000000000000000000000
16. 8 559 1 US-08-313 Sequence 14, Applife 6 76 4 5827390-10 Patent No. 542739016.6 76 4 5733901-11 Patent No. 542739016.6 76 4 5733901-11 Patent No. 542739016.3 76 2 US-08-36-35 Sequence 55, Applife 3 77 2 US-08-634 Sequence 55, Applife 3 77 2 US-08-634 Sequence 55, Applife 3 77 2 US-08-646 Sequence 55, Applife 3 77 2 US-08-465 Sequence 55, Applife 1 1045 2 US-08-465 Sequence 55, Applife 1 1045 2 US-08-465 Sequence 55, Applife 1 1045 2 US-08-465 Sequence 3, Applife 1 1045 2 US-08-467 Sequence 3, Applife 1 1045 2 US-08-53- Sequence 4, Applife 1 105-105 2 US-08-547 Sequence 4, Applife 1 105-105 2 US-08-547 Sequence 4, Applife 1 105-105 2 US-08-547 Sequence 4, Applife 1 105-105 2 US-08-646 Sequence 6, Applife 1 105-105 2 US-08-646 3 Sequence 6, Applife	955	atch cal Si 60 1.0RQY LQRQY	LENGE TOPEC MODER TITLE TITLE TITLE TITLE TOMPUT COMPU	ce 2	-955-	74455
1 US-08-313 Sequence 14, APP114 5482709-10 Patent No. 5482709-10 4 5273901-11 Patent No. 5482709-10 4 5273901-11 Patent No. 5273901 4 5273901-11 US-08-35-325 Sequence 55, App112 US-08-486 Sequence 65, App114 US-08-487 Sequence 4, App114 US-08-487 Sequence 4, App114 US-08-487 Sequence 4, App114 US-08-487 Sequence 4, App115 US-08-634 Sequence 4, App116 US-08-634 Sequence 53, App117 US-08-634 Sequence 63, App117 US-08-634 Sequence 63, App118 US-08-634 Sequence 63, App18 US-08-634 Sequence 63,	905A-2	milarit ; Cons ; Cons ; Cons ; Coccogro	SEO HEOHOO PAHERONNES	Appli	05 A -	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
1 US-08-313 Sequence 14, Appl. 4 5273901-11 Patent No. 5273901 4 5273901-11 Patent No. 5273901 5 US-08-933 Sequence 47, Appl. 2 US-08-486 Sequence 55, Appl. 2 US-08-461 Sequence 55, Appl. 2 US-08-465 Sequence 55, Appl. 2 US-08-467 Sequence 55, Appl. 2 US-08-53 Sequence 29, Appl. 2 US-08-540 Sequence 4, Appl. 2 US-08-541 Sequence 4, Appl. 2 US-08-541 Sequence 4, Appl. 2 US-08-634 Sequence 63, Appl. 3 PECOMBINANT 47 AND 31 kD COCOA PRECOMBINANT 47 AND 31 kD COCOA PRECOMSOR 2 ROBERTS 3 MW; 1529150 CN; 3 MG; 2: 3 COCOCORKCWEQYKEQERGEHENYHNHKKNR 11111111111111111111111111111111111	s	100. 100. rvati EQQQQ		atio	STA	559 76 76 77 77 77 77 77 77 77 77 77 77 77
0.8-313 - Sequence 14, Appli- 0.8-33 - Sequence 55, Appli- 0.8-480 - Sequence 55, Appli- 0.8-481 - Sequence 55, Appli- 0.8-481 - Sequence 55, Appli- 0.8-486 - Sequence 55, Appli- 0.8-461 - Sequence 55, Appli- 0.8-53 - Sequence 29, Appli- 0.8-54 - Sequence 4, Appli- 0.8-54 - Sequence 4, Appli- 0.8-634 - Sequence 63, Appl- 0.8-634 -	TAN	DRE	DRM RIPPY COPY COPY COPY COPY COPY COPY COPY C	υ _έ ,	ND	
10 Patent No. 5482700-10 Patent No. 5273901 131- Sequence 14, Appl. 132- Sequence 57, Appl. 133- Sequence 55, Appl. 134- Sequence 55, Appl. 136- Sequence 55, Appl. 136- Sequence 55, Appl. 137- Sequence 55, Appl. 138- Sequence 55, Appl. 139- Sequence 55, Appl. 139- Sequence 55, Appl. 139- Sequence 4, Appl. 139- Sequence 4, Appl. 139- Sequence 4, Appl. 139- Sequence 4, Appl. 139- Sequence 6, Appl. 139- Sequence 6, Appl. 139- Sequence 6, Appl. 140- Sequence 6, Appl. 151- Sequence 6, Appl. 152- Sequence 6, Appl. 153- Sequence 6, Appl. 154- Sequence 6, Appl. 155- Sequence 6, Appl. 155- Sequence 6, Appl. 156- Sequence 6, Appl. 157- Sequence 6, Appl. 158- Sequence 6, Appl. 159- Appl. 159- Sequence 6, Appl. 159	DARD;	Scor Pred 0; 0 2000CQR 111111	/07955 COMBINECURSC 8 disk patibl patibl -DOS/N elease A: US/07/ 1993 : 2: Ids	9559	ALIGNM	73908
atent No. 527390: atent No. 53 Appl: aequence 55, Appl: aequence 65, Appl: aequence 4, Appl: aequence 4, Appl: aequence 4, Appl: aequence 4, Appl: aequence 63, Appl: aequenc	PR	471; NO. Misma CWEQY 	, . O * * * * * * * * * * * * * * * * * *		NTS	111 111 111 111 111 111 111 111 111 11
14, 5482700 54827000 55, Appli 55, Appli 55, Appli 55, Appli 55, Appli 55, Appli 4, Appli 4, Appli 9, Appli 19, Appli 63, Appli 63, Appli 63, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 11, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 16, Appli 16, Appli 17, Appli 18, Appli 19, Appli 19, Appli 19, Appli 19, Appli 19, Appli 19, Appli 10, Appli 10, Appli 11, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 16, Appli 16, Appli 17, Appli 17, Appli 18, Appl		m — m О·	·			Patterson Part Passerson Pasterson P
14, 5482700 54827000 55, Appli 55, Appli 55, Appli 55, Appli 55, Appli 55, Appli 4, Appli 4, Appli 9, Appli 19, Appli 63, Appli 63, Appli 63, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 11, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 16, Appli 16, Appli 17, Appli 18, Appli 19, Appli 19, Appli 19, Appli 19, Appli 19, Appli 19, Appli 10, Appli 10, Appli 11, Appli 11, Appli 11, Appli 11, Appli 11, Appli 12, Appli 13, Appli 14, Appli 15, Appli 16, Appli 16, Appli 16, Appli 16, Appli 17, Appli 17, Appli 18, Appl	66	1; e-36; s 0 RGEHE 	sion		66	ent No lence lence lence lence lence lence lence lence lence lence lence lence lence lence lence
273901. ### Applica A		3-3 " 2	, , , , , , , , , , , , , , , , , , , 			00000000000000000000000000000000000000
H H A B A B A B A B A B A B A B A B A B		EER C	מ מי			
GG		EGC ;	EINS			
R 140		1PS 0 R 140 R 140	•			

:-:

+ 262+

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 10:45:58 2000; MasPar time 7.09 Seconds 398.931 Million cell updates/sec

Tabular output not generated.

Description: Perfect Score:

Title:

Sequence:

\(\sum_{\text{0}}\) \(\sum

1 LQRQYQQCQGRCQEQQQGQR.....HENYHNHKKNRSEEEEGQQR 60

Scoring table: Searched: PAM 150 Gap 11 142080 segs, 47172406 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

pir62 1:pir1 2:pir2 3:pir3 4:pir4

Database:

Statistics: Mean 32.973; Variance 71.563; scale 0.461

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

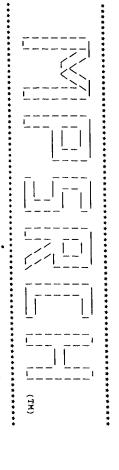
SUMMARIES

No.	Score	Query Match	Length	D B	ID	Description	Pred. No
1	471	100.0	566	2	S22477	vicilin precursor - c	8.50e-6
2	188	39.9	605	N	w	-globulin type	.81e-
w	181	•	509	2	S08059	lobulin type	. 30
	168	35.7	588	—	FWCNAB	lobulin	1.84e-1
ر.	125	26.5	1038	N	T02634	p protein homol	Ģ
6	121	25.7	613	2	S27770	-	4.82e-0
7	116		564	N	S37241	н.	2.69e-0
œ	114		495	,	I37062	rin s	5.32e-0
9	114		605	سر	137061	3	
10	114		635	ب	137060	involucrin L - gorill	5.32e-0
11	114		835	ب	157441	involucrin - oranguta	5.32e-05
12	111	23.6	522	س	137037	involucrin - common g	1.46e-04
13	111	·	551	2	5194	prunin 1 precursor -	1.46e-04
14	111	·	708	N	A53185	G-box-binding factor	1.46e-04
15	110	ω.	1390	N	T14004	trfA protein - slime	2.05e-04
16	109	ω.	1094	2	S49313	protein kinase – slim	2.86e-04
17	107	22.7	1457	N	T14577	protein kinase YakA (5.57e-04
18	105	22.3	285	ب	146207	$^{\circ}$	1.08e-03
19	105-	22.3	905	۳	S	regulatory protein SN	1.08e-03
20	104	22.1	139	N	A26892		1.50e-03
21	104	22.1	1655	N	φ		1.50e-03
22	103	21.9	289	N	S74632	thetical pro	2.08e-03
23	103	21.9	395	2	543344	sex-determining prote	2.08e-03

30-Sep-1993

4 4 U	. Δ . ω	42	41	40	39	3 8	37	36	3 5	34	ω ω	32	31	30	29	28	27	26	25	24
95	9 9	96	97	97	97	98	98	98	98	98	99	99	99	100	100	100	100	100	101	103
20.2	20.4	20.4	20.6	20.6	20.6	20.8	20.8	20.8	20.8	20.8	•	21.0	21.0	21.2	21.2	21.2	21.2	21.2	21.4	21.9
338	1021	398	2359	625	544	1010	560	517	507	203	1761	930	758	2150	1898	637	255	47	1154	395
2	งผ	N	2	N	μ.	N	۳	_	N	N	N	N	N	N	_	N	N	2	N	2
S04321	S33791	S21883	T03094	A34615	136911	T13167	A40125	FWPMLA	T06452	S09559	T13675	T08588	S54522	S71629	A45973	S35221	A60637	JC5557	S69206	S35565
legumin B (clone pCD3	nding protein	DNA-binding protein P	A-kinase anchor prote	profilaggrin - rat (f	involucrin L - dourou	Lola-like protein - f	involucrin - pygmy ch	legumin A precursor -	probable legumin A pr	legumin A (clone pCD4	hypothetical protein	hypothetical protein	hypothetical protein	sensory transduction	trichohyalin - human	globulin Begl precurs	merozoite antigen LPM	arginine/glutamate-ri	-	sex-determining prote
2.72e-02	1.98e-02	1.98e-02	1.44e-02	1.44e-02	1.44e-02	1.05e-02	1.05e-02	1.05e-02	1.05e-02	1.05e-02	7.60e-03	7.60e-03	7.60e-03	5.51e-03	5.51e-03	5.5le-03	5.51e-03	5.5le-03		2.08e-03

Db 81 LOROYOGCO Oy 81 LOROYOGCO OY 81 LOROYOGCO RESULT 2 ENTRY 2 ENTRY 30 ALTIERNATE_NAMES se ORGANISM 4£ DATE 31	Query Match Best Local Similarity 100.0%; Matches 60; Conservative	1 - 24 25 - 566 UMMARY #	S ICATION S	##molecule_type mRNA ##residues 1-45 ##rcross-references E	##molecule_type DNA ##residues 1-5 ##cross-references ##cross-references	*cross-references MUID:92284309 *accession S22477	400000000000000000000000000000000000000	journal		RESULT 1 ENTRY V TITLE V ORGANISM #:
LCROYOGCGGRCCEOQGGREQQCCORKCWEQYKEQERGEHENYHNHKKNRSEEEEGQQR 140	100.0%; Score 471; DB 2; Length 566; Similarity 100.0%; Pred. No. 8.50e-68; 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	#domain signal sequence #status predicted #label SIG\ #product vicilin #status predicted #label MAT length 566 #molecular-weight 65513 #checksum 3693	211/1; 269/3; 296/3; 391/3; 502/1 #superfamily glycinin seed; storage protein	##molecule_type mRNA ##rosidues 1-452 ##jabel MC2 ##rcross-references EMBL:X62626	##molecule_type DNA ##residues 1-566-##label MCH ##cros-references EMBL:X62625	ICES MUID:9228@309 S22477	vicilin genes of cocoa and cotton raise questions about vicilin evolution.		S22477; S22478; S18105; S22050 S22477	S22477 *type complete v1c11in precursor - cacao *formal_name Theobroma cacao *common_name cacao 30-Sep-1933 *sequence_revision 30-Sep-1993 *text_change 17-Mar-1999



Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 10:41:27 2000; MasPar time 239.76 Seconds 17.351 Million cell updates/sec

Tabular output not generated

Run on:

Description: Perfect Score: Sequence: 2505-09-331-631-7 -(81-140) from 0509331631.pep 471 (3 of 3)

LOROYOCCOGRECOEOOOGOR.....

HENYHNHKKNRSEEEEGQQR 60

150MD10351

PAM 150 Gap 11 225878 segs, 69334122 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Searched:

Scoring table:

Database: sptrembl12

1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle 9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified 13:sp_vertebrate 14:sp_virus

Statistics: Mean 32.878; Variance 66.854; scale 0.492

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description	Pred. No.
. :	471	100.0	525	5	043358	VICILIN PRECURSOR.	2.19e-72
2	133	28.2	411	u	P91419	CODED FOR BY C. ELEGAN	1.92e-08
w	125	26.5	1038	s	060983	REP-LIKE.	3.90e-07
•	119	25.3	171	11	061118	GLUTAMINE REPEAT PROTE	3.59e-06
v	119	25.3	018	10	Q9ZWI3	PV100.	3.59e-06
6	116	24.6	564	10	Q43673	LEGUMIN, LEGUMIN-RELAT	1.08e-05
7	116	24.6	905	v	018260	ZC101.1 PROTEIN.	1.08e-05
80	114	24.2	495	σ	Q28425	INVOLUCRIN.	2.22e-05
9	114	24.2	605	6	028424	INVOLUCRIN.	2.22e-05
10	114	24.2	635	6	Q28423	INVOLUCRIN.	2.22e-05
11	112	23.8	838	տ	Q9Y0C9	RAS INTERACTING PROTEI	4.5/e-U5
12	111	23.6	551	10	Q43607	PRUNIN PRECURSOR.	6.54e-05
. 13	110	23.4	304	£.	015409	CAGH44 (FRAGMENT).	9.36e-05
14	110	23.4	388	Ġ	016500	C03A7.4 PROTEIN.	9.366-05
15	110	23.4	388	Ç	016501	C03A7.7 PROTEIN.	9.350-05
16	110	23.4	1390	v	077033	TRFA.	9.36e-US
17	109	23.1	1094	v	Q23915		1.340-04
18	107	22.7	1457	5	044011	PROTEIN KINASE YAKA	2. /2e-04
19	104	22.1	139	11	Q62006	OPA REPEAT (FRAGMENT).	7.81e-04
20	104	22.1	406	N	087306	ELPB2.	/. ale. 04

5	A (<u>د</u> د	A .	4	40	39	38	37	36	u G	ω 4	ω ω	32	31	30	29	28	27	26	25	24	23	22	21	
100	100	2 5	101	101	101	101	101	101	102	102	102	102	102	102	102	103	103	103	104	104	104	104	104	104	
21.2	21.2	21.7	21.4	21.4	21.4	21.4	21.4	21.4	21.7	21.7	21.7	21.7	21.7	21.7	21.7	21.9	21.9	21.9	22.1	22.1	22.1	22.1	22.1	22.1	
2150	637	234	3726	2473	2462	2253	624	399	418	416	388	378	378	218	57	438	420	289	2074	2023	1655	652	622	445	
G,	0		_	1	11	11	u	11	11	11	σ	N	N	N	4	տ	11	N	11	4	ഗ	4	v	G	
Q23863	003678	035859	061329	Q9WU42	Q9WUC1	Q9WU43	094486	035863	035862	035957	044606	031328	008264	044782	Q9Y4J1	016502	035861	P72769	088542	075557	Q24754	015410	Q94918	016511	
HISTIDINE KINASE A.		SEX DETERMINING PROTEI	AT MOTIF-BINDING FACTO	SILENCING MEDIATOR OF	SILENCING MEDIATOR OF	MEDIATOR	0	SEX DETERMINING PROTEI	DETERMINING	DETERMINING	R09B5.5 PROTEIN.	ERPJ.	ERPB2.	ERPB.	SCAl PROTEIN (FRAGMENT	7			OPA-CONTAINING PROTEIN			CAGH45.	VEIN PROTEIN PRECURSOR	CO3A7.14 PROTEIN.	
3.13e-0	3.13e-03	3.13e-03	2.22e-03	2.22e-03			2.22e	2.22e-03	1.5/e-03	1.57e	1.5/e	1.57e-03	1.5/e-03	1.5/e-03	1.5/e-03	1.11e-03	1.11e-03	1.11e-03	/. BIE-04	7.81e-04	7.Ble-04		٠	7.81e	

ALIGNMENTS

RESULT ID P AC P DT 0	Ov Db	S T T K R R R R R R R R R R R R R R R R R	ID SOCIOLO CON CONTROLO CONTRO
LT 2 P91419 PRELIMINARY: PRT: 411 AA. P91419; O1-MAY-1997 (TrEMBLrel. 03, Created) O1-MAY-1997 (TrEMBLrel. 03, Last sequence update)	Query Match 100.0%; Score 471; DB 10; Length 525; Best Local Similarity 100.0%; Pred. No. 2.19e-72; Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 81 LOROYOOCOGRCQEQQGGREQOOCORXCWEQYKEQERGEHENYHNHXKNRSEEEEGQOR 140 81 LOROYOOCOGRCQEQQGGRAGOOCCRXCWEQYKEQERGEHENYHNHXKNRSEEEEGQOR 140	SEQUENCE FROM N.A. TISSUE-LEAVES; MEDLINE; 92288309. "Comparison of the structure and nucleotide sequences of vicilin genes "Comparison of the structure and nucleotide sequences of vicilin genes "Comparison of the structure and nucleotide sequences of vicilin evolution."; plant Mol. Biol. 18:1173-1176(1992). PLANT MOL. BIOL. 18:1173-1176(1992). EMBL; X62625; CAA44493.1; EMBL; X62625; CAA44494.1; EMBL; X62625; CAA44494.1; EMBL; X62625; CAA44493.1; EMBL; X62625; CAA4493.1; EMBL; X62625; CAA44493.1; EMBL; X62625; CAA4493.1; EMBL; X62625; CAA44493.1; EMBL; X62625; CAA44493.	LT 1 Q43358 PRELIMINARY; PRT; 525 AA. Q43358; Q1-NOV-1996 (TrEMBLrel. 01, Created) Q1-NOV-1996 (TrEMBLrel. 01, Last sequence update) Q1-NOV-1999 (TrEMBLrel. 12, Last annotation update) VICILIN PRECURSOR. CSV. CSV. CSV. CSV. CSV. CSV. CSV. CSV

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Uxford Mulecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm Run on:

Sat May 13 10:57:55 2000: MasPar time 6.44 Seconds 292.770 Million cell updates/sec

Tabular output not generated.

Intle:

Sequence: 1 PEDPQRRYEECQQECRQOEERQOPOCQQRCLKREEQEEQQO 40

----::; table: PAM 150

Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 04

Listing first 45 summariles

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Mean 30.834; Variance 62.690; scale 0.492

Statistics:

Database:

pir62 1:pir1 2:pir2 3:pir3 4:pir4

SUMMARIES

2221117654321 2221087654321	Result No.
314 199 1100 1110 1110 1110 1110 1110 111	Score
98.1 47.5.2 34.4.4 31.9 31.9 31.9 31.9 31.9 31.9 31.9 31.9	Query Match
588 509 605 566 566 566 1403 1403 1403 1407 314 643 339 339 339 339 339 339 339 319 319 31	Length
	. BB
FWCNAB \$08059 \$08059 \$22438 \$22448 \$41448 \$41448 \$41448 \$41457	ID
alpha-globulin B prec alpha-globulin type B alpha-globulin type A vicilin precursor - c pros protein - fruit homeotic protein protein neuronal precursor pr hypothetical protein protein kinase (EC 2 protein kinase Yaka (transcription initiat transcription factor glutamine-rich protein - flola-like protein - frecu alpha/beta-gliadin pr gamma-hordein 1 precu Mopa box protein - mo homeotic protein BarH regulator protein *Ime apolipoprotein A-IV - KIAA0642 protein - human	Description
9.85e-43 6.00e-41 1.50e-21 1.60e-05 1.60e-	Pred. No.

5	44	4 3	42	41	40	39	38	37	3 6	35	34	ω u	32	ω L	30	29	28	27	26	25	24
86	86	86	86	86	96	86	86	86	87	87	87	87	88	88	88	88	88	89	89	89	89
26.9	26.9	26.9	26.9	26.9	26.9	26.9	26.9		27.2	27.2	27.2	27.2	27.5	27.5	27.5	27.5	27.5	27.8	27.8	27.8	27.8
5262	4957	1027	924	538	454	404	326	319	905	873	671	173	885	872	795	467	285	3190	450	296	296
N	N	N	N	N	~	N	N	N	,_	Ŋ	N	N	N	N	2	ш	۳	N	2	N	2
T03454	T03455	T13071	B41359	S29521	A46498	S13741	D22364	A22364	RGBYS5	B53225	T02504	A47303	T09225	T09226	T09227	A49377	146207	T13828	S42825	S07361	A27319
ALR protein - human	ALR protein - human	CLOCK protein - fruit	potassium channel pro	casein kinase I homol	glucocorticoid-sensit	regulatory protein OP	alpha/beta-gliadin pr	alpha/beta-gliadin pr	regulatory protein SN	ecdysone-induced prot	hypothetical protein	FTZ-F1 steroid recept	A kinase anchor prote	A kinase anchor prote	A kinase anchor prote	involucrin - mouse	involucrin - dog	CREB-binding protein	serum response factor	alpha/beta-gliadin pr	gliadin - wheat
7.20e-02	7.20e-02	7.20e-02	7.20e-02	7.20e-02	7.20e-02	7.20e-02	7.20e-02	7.20e-02	5.16e-02	5.16e-02	5.16e-02	5.16e-02	3.69e-02	3.69e-02	3.69e-02	3.69e-02	3.69e-02	2.63e-02	2.63e-02	2.63e-02	2.63e-02

Db 80 PEDPQRE 	Query Match Best Local Simi Matches 39;	1-25 26-588 417	##molecule_type ##residues COMMENT This is a: CLASSIFICATION #sup KEYWORDS glyce	#Journal #title #accession #astatus	#accession A30838 ##molecule_type mRNA ##residues 1-88 ##cross references G ##cxperimental_sourc S06398 REFERENCE S06398	ACCESSIONS REFERENCE #authors #journal #title	RESULT 1 ENTRY TITLE ALTERNATE_NAMES ORGANISM DATE
PEDPORRYEECOOECROOEERORPOCOORCLKRFEQEOOO 119	98.1%; Score 314; DB 1; Length 588; Similarity 97.5%; Pred. No. 9.85e-43; Gonservative 1; Mismatches 0; Indels 0; Gaps 0;	*domain signal sequence *status predicted *label SIG\ *product alpha-globulin storage proprotein *status predicted *label MAT\ *binding.site carbohydrate (Asn) (covalent) *status predicted predicted *molecular-weight 69729 *checksum 941	_type_DNA	t Mol. Biol. (1987) 9:533-546 Iopmental blochemistry of cottonseed embryogene mination. XIX. Sequences and genomic organizational plant (vicilin) genes of cottonseed. In the compared with conceptual translation	D m l	A30838; S06911 Chlan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L. Chlan, C.A.; Biol. (1986) 7:475-489 Plant Mol. Biol. (1986) 7:475-489 Developmental biochemistry of cottonseed embryogenesis and germination XVIII cDNA and amilno acid sequences of members of the storage protein families.	FWCNAB *type complete alpha-globulin B precursor (clone C72) - upland cotton seed storage protein; vicilin precursor *formal_name Gossypium hirsutum *common_name upland cotton 30-sep-1991 *sequence_revision 30-Sep-1991 *text_change 16-Jul-1999

**************** (MI)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 12:12:57 2000: MasPar time 59.01 Seconds 7.246 Million cell updates/sec

Tabular output not generated

Description: Perfect Score: Sequence: Scoring table: PAM 150 Gap 11 >US-09-331-631-23 -(1-33) from US09331631 pep 287 1 RSGRGECRRQCLRRHEGQPWETQECMRRCRRRG 33

131253 segs, 12956647 residues

Searched:

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: a-issued 1:5A_COMB 2:5B_COMB 3:PCT_COMB 4:backfiles1

Statistics: Mean 22.066; Variance 101.322; scale 0.218

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

23	22		20		18	11		1,		14	13	12	11	5		٥	œ	7	5	ر.	4	w	2	_		Result
69	70	70	70	70	70) <u> </u>	7.1	74	74	74	74	74	74	-	7 .	74	74	79	86	86	87	87	87	787	, ,	Score
24.0	24.4	24.4	24.4	24.4	24.4	24.) i		25.8			25.8	•) ())	25.8		27.5		30.0	30.3	30.3	30.3	100.0		Query Match
900	642	642	384	384		1 1	3005	195	195	195	188	188	PBA		J. J.	55	55	587	566	566	188	188	188			Length
2	_	N	2	~		, ,	4	N	ᆫ	N	_	2		2 •	_	N	N	_	۰.	-	N	-	· N	,	ا د	DB
US-08-813-	US-08-414-		80.			-09-73	234	us-08-569-	US-08-469-	0S-08-609-	US-08-469-	05-08-569-		HE . 08 - 1.09 -	٠	US-08-569-	US-08-609-	US-07-955-	US-07-955-	05-07-955-	9-80-	0 0	0	0 0	115-08-822-	ID
Sequence 4,	Sequence 45,		Sequence II.			17		Sequence 7,	Sequence /,	,		,	, ·	٦.	Sequence 3,	Sequence 3,	Sequence 3,	, ,	, ,) ,) ,) <u>+</u>	٠.		_	Sequence 1,	Description
Application	Applicact		777		~ `	Applicat	5223423.	Applicatio	100	or tracks	Appracace C	- }	Applicatio	Applicatio	Applicatio	Application	Appricació	Applicaci	700110011	74071070	100010000000000000000000000000000000000	100 100 to	Applicati	annlicati	Applicatio	
0.0		7 200+01	7 20e+01	7 20e+01	7.20e+01	7.20e+01	6. Uye+UI	0.00	3 670+01	3 670+01	3 676+01	3 670+01		3.67e+01		3.076.01	3 670+01	3 670+01		,			3.85e+00	3.85e+00	3.78e-17	Pred. No.

22222222222222222222222222222222222222	
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
2244.0 2244.0 2244.0 2244.0 2233.7 2233.0 2233.0 2233.0 2233.0 2233.0 2233.0 2233.0	
1101 1118 1118 1112 1122 1122 1122 1122	
2424444444444444	
US-08-469- US-08-838- US-07-934- US-08-838- US-08-978- US-08-656- US-08-656- US-08-670- US-08-484- US-08-4484-	
Sequence 2, Sequence 2, Sequence 2, Sequence 6, Sequence 7, Sequence 7, Sequence 1, Sequence 1, Sequence 1, Sequence 2, Sequence 2, Sequence 2, Sequence 2, Sequence 3, Sequence 3, Sequence 3, Sequence 3, Sequence 3, Sequence 4, Sequen	
Application Applic	
8.51e+01 8.51e+01 8.51e+01 8.51e+01 8.51e+01 1.01e+02 1.01e+02 1.01e+02 1.01e+02 1.04e+	

88888	ននននន	8888888	888888888	366666666666666666666666666666666666666	S XX X X X X X X X X X X X X X X X X X
TELECOMMUNICATION INFORMATION: TELEPHONE: (515) 248-4800 TELEPAX: (515) 248-4844 INFORMATION FOR SEQ ID NO: 1:	APPLICATION NUMBER: 05 07 700 FILING DATE: 05 JUN-1900 ATTORNEY/AGENT INFORMATION: NAME: Yates, Michael J. REGISTRATION NUMBER: 0124C		CITE: Des MOLHES STATE: IOWA STATE: TOWA COUNTRY: USA ZIP: 50309 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible COMPUTER: IBM PC COMPATIBLE COMPUTER: PATENTIN Release #1.0, Version #1.30 SOFTWARE: Patentin Release #1.0, Version #1.30	H W C C T T I I I I I I I I	1 S-08-822-561 xxxxx equence 1, A

```
ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     38888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local S
Matches 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 11, Application US/08569063C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  XXXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-569-063C-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        y Match 100.0%;
Local Similarity 100.0%;
hes 33; Conservative
                                                                      TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                                       APPLICATION NUMBER: US 08
FILING DATE: 01-MAR-1995
ATTORNEY_AGENT INFORMATION:
NAME: EVANS, JOSEPH D
REGISTRATION NUMBER: 26.2
                                                                                                                                                                                         APPLICATION NUMBER: US 08/469,427
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
JENCE 34 AA; 4188 MW; 3921 CN;
                                                                                                                  REFERENCE/DOCKET NUMBER: 10
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: PAJUSOLA, KARTI
TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
TITLE OF INVENTION: DNA CODING THEREFOR
NUMBER OF SEQUENCES: 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RSGRGECRROCLRRHEGOPWETQECMRRCRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RSGRGECRRQCLRRHEGOPWETQECMRRCRRRG
OLECULE TYPE:
YPOTHETICAL: N
AGINAL SOURCE:
                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0
                                                                                                                                                                                                                                                                                                                                                    STREET: 1200 C. CITY: Washington
                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/569,063C
                                          STRANDEDNESS:
                                                                                                         TELEPHONE:
                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                               LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS
                                                      amino acid
                                                                                                                                                                                                                                                                                                                             20005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino acid
                                                                                                                                                                                                                                                                                                                                                 R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Application US/08569063C
                                                                                                                                                                                                                                                                                                                                                                     E: Evenson, McKeown,
                                                                 188 amino acids
                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                            OLOFSSON, Birgitta
ALITALO, Kari
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ERIKSSON, Ulf
                                 linear
                                                                                                          (202)
                      protein
                                                                                                                                                                                                                                        06-DEC-1995
                                          single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD:
                                                                                                          628-8800
                                                                                                                                                                                    US 08/397,651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 287; DB 2; L
Pred. No. 3.78e-17;
0; Mismatches 0;
                                                                                                                                         26, 269
                                                                                       11:
                                                                                                                                  1064/41979CP3
                                                                                                                                                                                                                                                                                                                                                                      Edwards & Lenahan,
Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0
```

밁

137 SPRPLCPR-CTQHHQ-RP-DPRTCRCRCRRRS 165

```
RESULT
ID U
                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             န္
       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 11, Application US/08469427A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  XXXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-469-427A-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                137 SPRPLCPR-CTQHHQ-RP-DPRTCRCRCRRRS 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 11, Patent No. 5
                                                                                                                     TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 188 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYPOTHETICAL: NO ORIGINAL SOURCE: TISSUE TYPE:
                                                                                                                                                                                                                                                                          COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
OPETWARE: PATENTIA PELEASE #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,427A
                                                                                                                                                                                                                                             CLASSIFICATION: 435
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Eriksson, Birgitta
APPLICANT: Olofsson, Birgitta
APPLICANT: Alitalo, Kari
                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 41
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS:
                                                                     MOLECULE TYPE: PI
                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Alitaio, I
APPLICANT: Pajusola,
                                                                                                                                                                                                                                                                                                                                         STATE: L~
                                                                                                                                                                                                                                                                                                                                                                                                             ITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B
ITLE OF INVENTION: DNA CODING THEREFOR
UMBER OF SEQUENCES: 17
                                                                                                                                                                                                                                                                                                                                                           STREET: LL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE TYPE: Primar Librobatcoma
RCE 188 AA; 21261 MW; 161002 CN;
                                                                                                                                                                                            NAME: Evans, Joseph D
REGISTRATION NUMBER:
                                                                                                                                                                                                                          APPLICATION NUMBER: US 08/397,651 FILING DATE: 01-MAR-1995
                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                           TOPOLOGY:
                                                                                                   STRANDEDNESS:
                                                                                                               TYPE: amino acid
                                                                                                                                                             TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30.3%;
Similarity 40.6%;
13; Conservat'
Similarity
13; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION:
                                     UE TYPE: human fibrosarcoma
188 AA; 21261 MW; 161002 CN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              l, Application US/08469427A
5607918
Conservative
                                                                                                                                                                                                                                                                                                                                                                               E: Evenson, McKeown,
1200 G Street, N.W.,
                                                                                           linear
          30.3%;
                                                                                                                                                               (202)
                                                                                                                                                                                                                                                                 06-JUN-1995
                                                                               protein
                                                                                                   single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                               628-8800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 87; DB 2;
Pred. No. 3.85e+00
          Score 87; DB 1; 1
Pred. No. 3.85e+00
                                                                                                                                                                                             26,269
                                                                                                                                                                                    41979cp2
   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                Edwards & Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ω
W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          188
                     Length 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 188
   9
                                                                                                                                                                                                                                                                                                                                                                                          Lenahar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <u>ب</u>
    ω
••
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
    Gaps
    w
••
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            w
T.
```

******** (MT)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 12:54:20 2000: MasPar time 59.05 Seconds $6.143 \ \text{Million}$ cell updates/sec \ 550 \ D NO:28

Tabular output not generated

Description.

Perfect Score:
Sequence: Scoring table: PAM 150 Gap 15 1 LDPIRQQQLCQMRCQQQEKDPRQQQQCK 28 . US-09-331-631-28 1-281 from-8809331631.pep

131253 seqs, 12956647 residues

Searched:

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: Mean 19:522; Variance 73:773; Beale 0:265 a-issued
1:5A_COMB 2:5B_COMB 3:PCT_COMB 4:backfiles1

Statistics:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

522	φ.	•	0	9 6	8	7 6	6	Ф	6	13 6	60	7	10 7	۷ 7	7	7 7	-	-	4 104	3 10	2 11	1 11	Result No. Score
29	29.	29.	29.	29.	29.	29.	29.	29.	30.	30.	30.	32.	32.	2 32.0	32.	32.	34.	43.	46.	46.	52.	52	Query Query e Match
108	108	89	68	89	89	89	970	678	737	737	2441	3144	3144	3144	760	351	361	587	805	805	566	566	Length
		N			N			w		-	_	_	1	N	-	_	-	۳	-	-	-	-	DB
	8-4	US-08-486-	-08-4	US-08-461-	US-08-634-	US-08-465-	US-08-673-	I-US	US-08-185-	US-08-185-	-08-19	US-08-453-	US-08-246-	US-08-457-	8-19	US-08-468-	US-08-415-	US-07-955-	CT.	8-04	US-07-955-	Ç	ID
	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Seguence	Sequence	Description							
٤	24	48,	8	48 ,	48,	~	7,	ω `	,	4	,	,		42,	'n	16,	٠,	23,	'n	Ņ	N	2,	on
	Applicati	Applicati	Applicati	Applicati	Applicati		Applicatio		Applicatio		-	Applicatio	Applicatio	Applicati	Applicatio	Applicati	Applicatio	Applicati	Applicatio	Applicatio	ica	Applicatio	* 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	2.10e+01	•	٠	2.10e+01	٠	•		1.71e+01	1.39e+01	1.39e+01	1.13e+01	٠	6.04e+00			3.96e+00			5.45e-03			2.81e-04	Pred. No.

5	44	3	42	41	40	39 9	38 8	37	36	35	34	33	32	31	30	29	28	27	26	25	24	
62	62	62	63	63	63	65	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	
27.6					28.0									•	•	٠	29.3			•		
795	252	252	816	713	428	138	2703	604	604	360	303	171	171	171	171	171	108	108	108	108	108	
-ب	N	ب	2	_	-	N	_	2	N	N	H	N	N	N	N	N	N	N	N	N	2	
US-07-716-	US-08-974-	US-08-685-	US-08-267-	US-08-190-	US-08-190-	US-08-684-	-80	-88	-80-	- 80-	-80-		-80-	-80-	-80-	-08-4	-08-4	US-08-486-	•	US-08-486-	US-08-634-	
Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence							
'n	7,	7,	9	63,	29,	'n	19,	12,	12,	Ņ	'n	62	62	52	62	62	53,	24,	53,	24,	24,	
Applicatio	Applicatio	Applicatio	Applicatio	Applicati	Applicati	Applicatio	Applicati	Applicati	Applicati	Applicatio	Applicatio	Applicati	Applicati	Applicati	Applicati							
4.75e+01	4.75e+01		3.88e+01			2.58e+01						2.10e+01		2.10e+01	2.10e+01	2.10e+01	2.10e+01	2.10e+01	٠	٠	2.10e+01	

RESULT ID U XX AC ×	do do	ž m o	RESULT TO COO CO C
US-07-955-905A-22 STANDARD; PRT; 566 AA. xxxxxx	83 ROYOOCOGRCOEQOGOREQOOC 105	Ouery Match 52.0%; Score 117; DB 1; Length 566; Best Local Similarity 60.9%; Pred. No. 2.81e-04; Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;	US-07-955-905A-2 STANDARD; PRT; 566 AA. XXXXXX Sequence 2, Application US/07955905A PATCHER 1 IN FORMATION: PRECURSOR TITLE OF INVENTION: PRECURSOR TITLE OF INVENTION: PRECURSOR TITLE OF INVENTION: PRECURSOR COMPUTER READABLE FORNA': MEDIUM TYPE: Floppy disk COMPUTER IBM PC Compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patentin Release #1.0, Version**1.25 (EPO) CURRINT APPLICATION DATA: APPLICATION NUMBER: US/07/955,905A FILING DATE: 21-JAN-193 INFORMATION FOR SEG ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 566 amino acids TYPE: amino acids TYPE: amino acids TYPE: amino acids TYPE: protein SEQUENCE 566 AA; 65613 MW: 1529150 CN:

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 09:00:19 2000; MasPar time 78.64 Seconds 8.907 Million cell updates/sec

Description: perfect Score: Sequence: Tabular output not generated. >US-09-331-631-25 (1-23) from US09331631.pep 177

MMRARFPLLLLGLVFLASVSVSF 23

C 2500 10 HO: 15

swiss-prot38 1:swissprot

Post-processing: Minimum Match 0% Listing first 45 summaries

Searched:

83857 segs, 30454973 residues

Scoring table:

PAM 150 Gap 11

Statistics: Mean 33.887; Variance 71.622; scale 0.473

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1100 1100 1100 1100 1100 1100 1100 110	Result
177 138 138 138 136 120 1120 1111 1111 1111 1111 1111 1111	Score
100.0 78.0 78.0 78.0 76.8 64.4 45.8 45.3 11.5 44.6 44.6 43.5 43.5	Query Match
605 639 439 439 554 445 554 554 560 676 676 676 676 676 676 676 676 676 6	Length
	DB B
GLCA_SOYBN GLCX_SOYBN PHSB_PHAVU PHSB_PHAVU PHSB_PHAVU GLCB_SOYBN CVCA_PEA CVCB_PEA CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL CANA_CANGL VCLC_PEA VCL_VICFA VCLVICFA CP78_MAIZE VST2_HEVME Y752_BORBU AH11_ARAHY LDL1_XENLA LYSA_BPP2 SCRC_RAT YPHE_BACSU WCAO_ECOLI YBF3_YEAST VST2_HEVPA	ID
BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH PHASEOLIN, BETA-TYPE P PHASEOLIN, ALPHA-TYPE PHASEOLIN, ALPHA-TYPE BETA-CONGLYCININ, BETA CONVICILIN PRECURSOR. CONVICILIN PRECURSOR. CONVICILIN PRECURSOR. CANAVALIN PRECURSOR. VICILIN PRECURSOR.	Description
3.71e-14 6.39e-09 1.39e-09 1.39e-09 1.39e-05 9.18e-05 9.18e-05 9.18e-05 9.18e-05 9.18e-05 9.18e-05 9.18e-00 1.36e+00 1.3	Pred. No.

4 5	44	د س	42	41	4	5	ים	ا د. 80	37	36	35	34	5.5) t	ب د د	<u>د</u> 1	3 0	29	200) k	3 1	26	25		<u>ي</u>	
72	72	72	72	73	1	7.0	73	73	73	73	73	73		7 .	74	74	75	ò	ìò	7 .	76	76	//	j .	77	
40.7	40.7	40.7	40.7	41.2		3	41.2	41.2	41.2	41.2	41.2	41.2		1 1	41.8	41.8	42.4	4.4		3 .	42.9	42.9	٠. ن		۵. ن	
690	669	405	142	000	,	904	892	837	642	393	202	100		£03	412	173	9/9	2	100	95.5	214	213		2,40	660	
۲	ب	_	۰,		۱ د	<u>۔</u> ۔	<u>, .</u>	ب	-	_	٠,	-	٠,	نــ		μ.	-	٠,	٠,		<u>ب</u>	۰	٠,	_		
PERO_DROME	FPS1_YEAST	LPPB_HALLN	LCA_CAVPO		BOT C HCV TA	TORS ECOLI	LDL2_XENLA	XYNZ_CLOTM	YBBD_BACSU	THOUSE OFF	NOOT KHOCA	F315 ECOL1	11003 3130	NUSM GORGO	GLUP_BRUAB	PRI_MEDIK	AMULMOUSE	Protection in	TTRO ROVIN	MRAY CHLTR	GLP1_PHANI	TT4T HUMAN		VSTO HEVBU	VST2_HEVMY	
PERUALDAGE PRECURSON (GENCEROL OFTENSOR /	OUTER MEMBRANE ANTIOES	SUPER VENDENNE ANTIGEN	σ	GENOME POLYPROTEIN (CO	SENSOR PROTEIN TORS (E	LOW-DENSITY LIPOPROTEI	ENDO-1, 4 -BETA-XILANASE	HIPOIRELICAL 70.0 NO E	HIPOINGIFON TO 6 KD I	NACH CENTUROSCHICK - C	THE PREVENCENACE TO	PHOSPHATE STARVATION-I	NADH-UBIQUINONE OXIDOR	GLUCOSE/GALACTOSE TRAN	PATHOGENESIS RELATED A	TOTAL COLCANO SECTION OF THE PROPERTY OF THE P	DESTINATION OF STANK	CELL SURFACE ADHESION	PHOSPHO-N-ACETYLMURAMO	GERMIN-LINE PROTEIN FA	LENGTH COUNTY TO THE PARTY OF T		STRUCTURAL PROTEIN 2 P	STRUCTURAL PROTEIN & T	
	1 760+01	1 760+01	760+01	1.76e+01	1.34e+01	1.34e+U1	1.040	340+01	1 3/0+01	1 346+01	1 34e+01	1.34e+01	1.34e+01	1.010101	1.016701	1 010-01	016+01	7.67e+00	7.67e+00	5.78e+00	100	5 780+00	5.78e+00	4.300+00		350+00

SEQUE TISSUE MEDLII Sebas: "Comp. cooff. plant p	
Glyclne. [1] SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEDATION: SEBASTIANI F.L., Farrel L.B., Schuler M.A., Beachy R.N.; SEBASTIANI F.L., Farrel L.B., Schuler M.A., Beachy R.N.; SEDASTIANI F.L., Farrel L.B., Schuler M.A., Beachy R.N.; PHART M.S. SEDASTORAGE PROTEIN ACCUMULATES DURING SEED LING. CARBON AND ISTROGEN SOURCE FOR THE DEVELOPING SEEDLING. CARBON M	RESULT 1 ID GLCA_SOYBN STANDARD; PRT; 605 AA. AC P13916; DT 01-JAN-1990 (Rel. 13, Created) DT 01-JAN-1990 (Rel. 14, Last sequence update) DT 15-DEC-1998 (Rel. 37, Last annotation update) DT 15-DEC-1998 (Rel. 37, Last annotation update) DE BETA-CONGLYCININ, ALPHA CHAIN PRECURSOR. OS Glycine max (Soybean). OS Glycine max (Soybean). OC Glycine max (Soybean).

(MT)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 08:51:56 2000: MasPar time 6.17 Seconds 252.411 Million cell updates/scc

Tabular output not generated: 1 RSGRGECRROCLRHEGOPWETOECHRRCRRG 33 SQQ 10 ND. 73 |
PAM 150
Gap 11

Title:

Run on:

Description: Perfect Score: Sequence:

Post-processing: Minimum Match 0% Listing first 45 summaries

142080 segs, 47172406 residues

Searched:

Scoring table:

pir62 1:pir1 2:pir2 3:pir3 4:pir4

Database: Mean 30.542; Variance 60.292; scale 0.507

Statistics:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

20 22 23	16 17 18	113311100887654321	Result
79 78 77	79 80 80 80	287 927 927 927 927 927 927 927 927	Score
27.5 27.2 27.2 26.8	27.9		Query Match
484 482 45	428 112 525 91	55400 5588 5588 55400 5573	Length
NNNN	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	Ba
\$68952 \$65779 \$58208	5168 5157 5157 2167	A41822 C53234 B53234 FWCNAB FWCNAB C347059 C34707 T01662 T012662 T012662 T012662 T012662 T012662 T012662 T012662 T012662	ID
omega-hordothionin - gamma-thionin precurs protamine II-1 - pain	-8b - A wnt-8m mine 2	antimicrobial peptide astB/chuR-related pro globulin-10 - maize vicilin-like storage alpha-globulin B prec alpha-globulin type B ORF2 protein - Orf vi vicilin precursor - c globulin-l - maize vicilin-like storage globulin-like storage globulin-like storage neuron-glia adhesion BMP receptor precurso gene wnt8b protein -	Description
6.28 8.70	2.33e 3.24e 3.24e 4.52e 4.52e	7.27e-39 5.27e-03 7.50e-03 7.50e-03 7.14e-02 4.27e-02 4.27e-02 6.02e-02 6.02e-02 8.6e-02 1.66e-01 1.66e-01	Pred. No.

45	44	43	42	4		0	39	3 8	37		3 6	J.	34	L.) (بر د	3 1	č	, ,	<u>ي</u> د	28	27	0) i	25	24		
73	73	73	73	١,		74	74	74	/4	1 4	7 /	74	76	ò	7.	76	76	ò	1 .	76	76	77	-	1	77	77	•	
25.4	25.4	25.4	25.4	20.4	י מ	25.8	25.8	25.8	20.0		25 8	25.8	26.5		ν γ	26.5	26.5) i	26.5	26.5	26. B		3	26.8	20. a)	
4006	102	.92	0.2	, ,	л œ	1306	1073	Tag			47	45	707		103	102	7.5	: ;	7	47	4.	49/0		605	187	U	n O	
2	N	~		۰,	J	N	2) +	_	N.	N	_	, ,	J	N	٨		J	N	2		, 1	N	N		J	
T09070	S33336	513132	010/04	01010	S10755	T13592	556220	10000	10100	Calinon	F58208	C58208	20001		22223	S33332	C/TOOT		B75130	E58208	869810	TELOU	1	S06398	G64323	200	3 KB 3 O B	
probable tenascin A	protamine P2 - inesus	7	THE COMMENT OF THE COMMENT	•	protamine St2b - hors	נכר	propose memorane pro	Addition of the party of the pa	wascular endothelial	sperm histone P2 prec	protamine II-5 - pain	,	11	, 1	protamine P2 - gorill	procamine rz - pywy		nichable ribosomal pr	isu ribosomal protein	protamine II-4 - Pain	nypochector process	Contract Cal State Cal	portide evertherage ho	alpha-grobutin type A		2012 61	protamine I-1 - paint	
0.146.00	3 140 00	3 146+00	3 140+00	3.14e+00	3.14e+UU	3 4 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 399+00	2 296+00	2.29e+00	2.29e+00	2.236+00	3 300 00	2 246+00	1.20e+00	1.208-00	1 200 100	1 200+00	1.20e+00	1.20e+00	F. 100	1 308100	1 206+00	8.70e-01	0.700	8 702-01	8 700-01	8.70e-01	

ALIGNMENTS

•						
#accession H722 ##status ##molecule_type .*	#journal #title #cross-refere	DATE ACCESSIONS REFERENCE #authors	Qy 1 RSGRGE RESULT 2 RESULT 2 RITLE: COCCANICA	Query Match Best Local Sim Matches 33; Db 1 RSGRGE	REFERENCE **authors **submission **accession **status **molecule **residues	RESULT 1 ENTRY TITLE ORGANISM DATE ACCESSIONS
H72266 preliminary _type DNA	Pratt, N.S.; Phillips, C.A.; Pratt, M.S.; Phillips, C.A.; elberg, J.; Sutton, G.G.; Fleisberg, S.L.; Smith, H.O.; Venter (1999) 399:323-329 (e. for lateral gene transfer be eria from genome sequence of Tr ID:99287316	11-Jun-1999 *sequence_revision ll-Jun-1999 *text_change 11-Jun-1999 #17266 #172200 #17	RSGRGECRROCLRRHEGOPWETQECMRRCRRG 33 R 3	tch 100.0%; Score 287; DB 2; Length 33; al Similarity 100.0%; Pred. No. 7.27e-39; Indels 0; Gaps 0; 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0; RSGRGECRROCLRRHEGOPWETQECMRRCRRRG 33	##1022 Duvick_ J.P.; Rood, T.; Rao, A.G.; Marshak, D.R. submitted to the Protein Sequence Database, May 1992 ### A41822 ### preliminary _type protein _type protein 1-33 ##!abel DUV #length 33 #molecular-weight 4131 *checksum 3118	A41822 #type complete antimicrobial peptide - maize #formal_name Zea mays #common_name maize 28-May-1992 #sequence_revision 28-May-1992 #text_change 30-Sep-1993 A41822

しいト くいト トト・トひだ

taye t

(MT)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 08:35:40 2000; MasPar time 73.66 Seconds 24.806 Million cell updates/sec 102\ 2505.09-331-631-21 2000 Par 1000 Par 100

Tabular output not generated

Description: Perfect Score: Sequence: >US-09-331-631-21 (32-91) - from USU9331631.pep

TENPCAQRCLQSCQQEPDDL.....DTGATNQRHPPGERTRGRQP 60

Scoring table: PAM 150 Gap 11

Searched: 83857 segs, 30454973 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

swiss-prot38 1:swissprot

Database:

Statistics: Mean 33.775; Variance 51.458; scale 0.656

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Result
453 1296 1297 1199 93 93 75 75 77 74 74 74 74 74 74 77 77 77	Score
100.0 26.3 26.3 20.5 116.8 116.8 116.8 116.3 116.3 116.3 116.3 116.3 116.3 116.3 116.3 116.3 116.3 116.3	Query Match
614 626 639 639 6605 676 676 675 121 121 147 221 147 221 147 229 320 488 588 588 588 588 588 588 588 588 588	Query Match Length
بير نيز	ВВ
AHIL_ARAHY AHIZ_ARAHY GLCX_SOYBN GLCA_SOYBN GLCA_SOYBN KALM_HUMAN KALM_HUMAN KALM_GOSHI RPP_CAEEL VCLA_GOSHI RPP_CONTO YIZI_BURCE V70K_EPMY IPZT_SOLTU VHEL_PVMR YBJB_ECOLI CRUA_BRANA VCLB_GOSHI STC_DROME ETVL_MOUSE ETVL_MOUSE AGRI_CAEEL AGRI_	ID
ALLERGEN ARA H 1, CLON ALLERGEN ARA H 1, CLON BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH BETA-CONGLYCININ, ALPH BETA-CONGLYCININ PROTEIN C72-A PRECURS FRIBONUCLEASE PANCREATI INSERTION ELEMENT IS40 TO ROTEIN FROTEIN FROTEIN FRECURSOR (CRUCIFERIN PRECURSOR (VICILIN C72 PRECURSOR SHUTTLE CRAFT PROTEIN ERB1 PROTEIN (ETS TRAN HYPOTHETICAL 85.1 KD N AGRIN PRECURSOR. PROTEINASE INHIBITOR T WOUND-INDUCED PROTEINAS FROM PROTEIN FROM SHUTTLE CRAFT PROTEIN FROTEIN FROTEIN FROTEIN FROM PRECURSOR. PROTEINASE INHIBITOR T WOUND-INDUCED PROTEINASE INHIBITOR T	Description
1.78e-92 1.15e-09 1.15e-09 2.14e-01 4.64e-01 4.64e-01 4.64e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01 6.40e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01 6.80e-01	Pred. No.

444	44	4 3 3 0 9 8	35 37	33 43 2	29 30	26 27 28	254
6888	000	5 5 5 8 8 8	0 0 0 0 0 0	70 69	70 70	71	71 71
15.0	15.0	15.0 15.0	15.2 15.2	15.5 15.2	15.5 15.5	15.7 15.7 15.7	15.7 15.7
642	641	104 291 292	425 829 829	655 1382 148	106 147 275	1148 1148 1148	892 1148
	ر سو سو د		سر سو سر		ں ہر		سر س
HEMA_INCGL	CGD1_RAT	CYC_MACMA CGD1_XENLA CGD1_CHICK	FTSA_HAEIN VIRA_AGRT9 VIRA_AGRT6	HEMA_INCJH INSR_HUMAN VEGH_ORFN7	GAS4_ARATH IP2Y_SOLTU VGLM_PUUMB	AGTW_bandh AGTW_bandh AGTW_bandh	NIRA_EMENI VGLM_POUMK
HEMAGGLUTININ PRECURSO HEMAGGLUTININ PRECURSO		CYTOCHROME C. G1/S-SPECIFIC CYCLIN D G1/S-SPECIFIC CYCLIN D	CELL DIVISION PROTEIN WIDE HOST RANGE VIRA P	HEMAGGLUTININ PRECURSO INSULIN RECEPTOR PRECU VASCULAR ENDOTHELIAL G	GIBBERELLIN-REGULATED PROTEINASE INHIBITOR T M POLYPROTEIN [CONTAIN	M POLYPROTEIN PRECURSO M POLYPROTEIN PRECURSO M POLYPROTEIN PRECURSO	~
6.20e+00 6.20e+00	6.20e+00	6.20e+00 6.20e+00 6.20e+00	4.33e+00 4.33e+00 4.33e+00	3.01e+00 3.01e+00 4.33e+00	3.01e+00 3.01e+00 3.01e+00	2.09e+00 2.09e+00 2.09e+00	

oy Db	3 W O	SQ	무 무	388	388	888	888	გგ	C R	RT	₽ 5	건 건 건 건	R C	388	888	BB	D F	ដូ	AC ID	RE	
32 TENPCAQRCLOSCOQEPDDLKQKACESRCTKLEYDPRCVYDTGATNORHPPGERTRGROP 91	Query Match 100.0%; Score 453; DB 1; Length 614; Best Local Similarity 100.0%; Pred. No. 1.78e-92; Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Allergen. SEQUENCE 614 AA; 70283 MW; 1DDACF217EEC5F31 CRC64;	HSSP; P50477; 1CAW. PFAM; PF00546; Seedstore_7s; 1.	FMRI. 1.3883. AAA60346].	modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/	use by non-profit institutions as long as its content is in no way	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation.	CONVICILIN, CONGLYCININ, ETC.).	J. Clin. Invest. 96:1715-1721(1995). TI- SIMILARITY: TO 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,	<pre>"Recombinant peanut allergen ara n i expression and igE binding in patients with peanut hypersensitivity.";</pre>	Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;	SEQUENCE FROM N.A. STRAIN-CV. FLORUNER;	Arachis.	core endicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;		ALLERGEN ARA H 1, CLONE P17 (ARA H I)	01-NOV-1997 (Ref. 32, Last annotation update)	(Rel.	ST		ALLGUMENTO

RESULT N

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by,Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Tabular output not generated Sat May 13 09:04:49 2000; MasPar time 4.78 Seconds 227.008 Million cell updates/sec

505-09-331-631-25 (1-23) from USD9331-631.pep

55010 HO: 25

Scoring table: PAM 150 Gap 11 Description: C Perfect Score: Sequence:

1 MMRARFPLLLLGLVFLASVSVSF 23

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database:

Statistics: pir62
1:pir1 2:pir2 3:pir3 4:pir4 Mean 33.374; Variance 80.850; scale 0.413

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

23	>	21	20	19	18	17	16	15	14	13	12	11	10.	9	æ	7	σ	v	4	w	N	ш	NO. S
800	BO	18	81	82	82	82	83	84	94	94	105	111	111	114	120	136	138	138	138	170	177	177	Score
		•	45. 8	•				47.5	۳.	53.1			62.7	•	7	CD.	78.0					100.0	Query Match
7 1 1	141	909	502	659	400	286	. 438	547	463	463	459	445	445				436	430	421	639	0	605	Length
N	J	بس و	N	<u>_</u>	N	N	N	-	N	N	2	2	N	N	N	,	N	N	N	N	N	ь.	DB
H72064	755055	QRXLL1	G70193	B44212	F72350	7237	S35757	S51475	S06309	A27288	S08505	JQ2264	S00281	0657	0056	FWSYCB	A23498	S10156	A24810	B24810	S20007	FWSYBA	ID
metal transport prote	P TOTO :		a	structural protein 2	O	sugar ABC transporter	7K - 0	chro	precurso	vicilin precursor - f	vicilin - garden pea	canavalin - jack bean	canavalin - sword bea		convicilin precursor		phaseolin type alpha	alpha-phaseolin precu	phaseolin beta chain	beta-conglycinin alph	lycinin	beta-conglycinin alph	Description
7.49e+00	7 496+00	5.77e+00	5.77e+00	4.44e+00	4.44e+00	4.44e+00	3.41e+00	2.61e+00		1.68e-01	7.03e-03	1.18e-03	1.18e-03			.87e-		.56e-	2.56e-	6.41e-	5.92e-	5.92e-	Pred. No.

5	44	43	42	2	0	39	3 8	37	36	35	34	w w	32	31	30	29	28	27	26	25	24
76	76	76	76	77	77	77	77	77	77	78	78	78	78	78	78	78	78	79	79	79	80
42.9	42.9	42.9	42.9	43.5	43.5	43.5	43.5	43.5	43.5	44.1	44.1	44.1	44.1	44.1	44.1	44.1	44.1	44.6	44.6	44.6	45.2
696	534	381	336	686	660	433	271	246	124	734	734	732	704	685	684	684	492	405	402	67	449
N	۳	N	N	2	_	N	2	N	N	2	2	N	N	N	N	N	N	N	N	N	N
T13767	S75101	T13701	D71474	T13680	VHWWH2	G70345	T11692	E64627	S45788	T13785	T12602	T14233	T13503	T12129	T13491	T12151	T02458	G64971	T13499	в69936	S16319
NADH dehydrogenase -	hypothetical protein	NADH dehydrogenase su	probable muramoyl-pen	NADH dehydrogenase -	structural protein 2	hypothetical protein	NADH dehydrogenase su	hypothetical protein	probable membrane pro	NADH dehydrogenase -	NADH dehydrogenase -	NADH dehydrogenase -	NADH dehydrogenase F	NADH dehydrogenase -	NADH dehydrogenase -	NADH dehydrogenase -	hypothetical protein	hypothetical protein	NADH dehydrogenase su	hypothetical protein	secretin receptor - r
2.09e+01	2.09e+01	2.09e+01	2.09e+01	1.62e+01	1.62e+01	1.62e+01	1.62e+01	1.62e+01	1.62e+01	1.25e+01	9.70e+00	9.70e+00	9.70e+00	7.49e+00							

1 : 0 : 1

ALIGNMENTS

g 1 MMRAREPLLLLGLVFLASVSVSF 23

	8.28e-16 4.26e-15 6.41e-15 9.64e-15 9.64e-15 9.64e-15 9.64e-15 1.45e-14 2.18e-14 2.18e-14 3.27e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14 7.35e-14	IN-1 (FRAGMENT).	GLOBULIN-1 (GLOBULIN-1 (GLOBUL	QQQ1865 QQQ1865 QQQ1865 QQQ1257 QQQ1155 QQQ1155 QQQ1155 QQQ1256 QQQ1256 QQQ1256 QQQ1255 QQQ1256 QQQ1255 QQQ1256 QQQ1256 QQQ1255 QQQ1256 QQQ125	122 10 238 10 238 10 238 10 238 10 240 10 240 10 240 10 242 10 242 10 242 10 243 10 243 10 244 10 245 10 247 10 247 10	200 - 100 -	174 174 174 179 169 168 168 168 168 166 166 166 166 166 163 163 163 163	200
•	ed.		Description	SUMMARIES		ery tch L	Score	Result
	e a printed,	by chance to have he result being p distribution.	redicted core of t tal score	results 1 to the of the t	number an or eq y analys	. is th eater t erived	red. N	טי אט עס
		0.496	55.241; scale	Variance 6	32.348;	Mean	tics:	Statis
	human p_organelle unclassified	4:3P_ hc 8:s 12:sp_	10 3:sp_fur mammal 7:sp 11:sp_roder	#p_bactonte for the first for	mbl12 p_arche p_inver p_phage sp_vert	sptre 1:a 5:s 9:s 13:	 	Database
			ies	0% 45 summarie	num Match .ng first	g: Minimum Listing	processing	Post-pı
			esidues	69334122 re	8 seqs,	22587	ŭ.	Searched
					150 11	PAM 1 Gap 1	g table:	Scoring
20:54	1t. man algorithm econds updates/sec 5 60 10 / GRGRG 55	Arch Unith, U.K. Ltd h-Waterma 35.80 Sec n cell up	ocompulity of xford rch, u Masp 16.1	ollins, B Unive ights by tabase : 4:54 200 24 24 24 CVQRCRQ	A John F. C (c) 1993-199 stribution r - protein da . May 13 08:5 enerated. -09-331-631 -85) from U6	ase 3 right otein otein not (3)	de d	MPsrch_p Run on: Run on: Tabular Tatle: Descript Perfect Sequence
•					 			

44444444444444444444444444444444444444
100 99 99 91 91 88 88 88 88 88 88 88 88 88 88 88 88 88
244 221 221 231 231 231 231 231 231 231 231
897 456 456 810 810 625 625 625 649 855 331 311 311 857 188 905 57 1182 57 1182 676 676 676 676 676 676 676 676 676 67
110 110 110 110 110 110 110 110 110 110
Q17336 Q53478 Q53478 Q53478 Q52473 Q43358 Q43358 Q46043 Q16043 Q16043 Q18260 Q1
LET 858. INNER CENTROMERE PROTE ORF2 PROTEIN. PV100. VICILIN PRECURSOR. FILAGGRIN (FRAGMENT). TRICHOHYALIN (FRAGMENT). TRICHOHYALIN (FRAGMENT). STELPALPHA. ALPHA-GLIADIN STORAGE ZC101.1 PROTEIN. TYROSINE KINASE ACTIVA PUTATIVE SPLICING FACTIVA PUTANSCRIPTION FACTOR P INTEGRIN BETA 3 SUBUNI KINAGE ACTIVA PROTEIN. FROMIN PRECURSOR. HYPOTHETICAL 62.4 KD P F53H1.1 PROTEIN. EXTRACELLULAR MATRIX P L-AFADIN. F43G9.10 PROTEIN. HYPOTHETICAL 61.1 KD P HYPOTHETICAL 61.1 KD P HYPOTHETICAL 61.7 KD P
1.00e-03 2.86e-03 4.42e-02 8.61e-02 1.20e-01 1.20e-01 1.20e-01 2.31e-01 2.31e-01 3.20e-01 3.20e-01 3.20e-01 3.20e-01 3.20e-01 3.20e-01 4.43e-01 4.43e-01 4.43e-01 6.11e-01 8.42e-01 1.16e-01

RESULT ID Q AC Q	Qy Db	W & O	SO DE RECOCCOCCE PER RECOCCOCCE PER RECOCCOCCE PER RECOCCOCCE PER RECOCCE PER	RESULT AC Q DT 00 DT 0
JLT 2 Q03863 Q03863;	31 DEDDR 31 DEDDR	Query Match Best Local Si Matches 55	Hordeum vulgare (Eukaryota; Viridi euphyllophytes; S Poaceae; Hordeum. [1] SEQUENCE FROM N. A MEDLINE; 93287988 HECK G. R., CHAMBE "Barley embryo gl chromosome mappin MOI. Gen. Genet	ULT 1 Q03678 Q03678; Q1-NOV-1996 (TrE Q1-NOV-1996 (TrE Q1-NOV-1999 (TrE EMBRYO GLOBULIN. BEG1 OR GBL1.
PRELIMINARY;	RGGHSLQQCVQRCRQEF RGGHSLQQCVQRCRQEF	/ Match 100.0%; Local Similarity 100.0%; hes 55; Conservative	re (Barley), ridiplantae; ridiplantae; sysermatop leum. N.A. 7988. 7988. 7988. 7988. 7988. 7988. 7099- 209- 209- 209- 209- 209- 209- 209-	PRELIMINARY; 6 (TrEMBLrel. 01, 6 (TrEMBLrel. 01, 9 (TrEMBLrel. 12, BULIN.
PRT;	RPRYSHARCY PRYSHARCY	Score 425; Pred. No. 0; Misma	streptophyta; Nyta; Magnolic Nyta; Magnolic Nyta; Magnolic Nyta; Magnolic Sene, Begl: c gene, Begl: c gene, Begl: c pulation of ex D18(1923) PROTEIN-PROCES E PROTEIN FRO SEED STORAG N, ETC.) S; 1. S; 1. S9BCEB36	PR: Crea1 Last Last
122 A	QECRDE QECRDE	44	blocessic CF	637) juence
AA.	DEDDRRGGHSLQQCVQRCRQERPRYSHARCVQECRDDQQQHGRHEQEEEQGRGRG 85 	DB 10; Length 637; .48e-64; ches 0; Indels 0; Gaps 0;	Hordeum vulgare (Barley), and Triticum aestlvum (Wheat). Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum. [1] SEQUENCE FROM N.A. MEDLINE; 93287988. MEDLINE; 101287988. MEDLINE; 1012879888. MEDLINE; 101287988. MEDLINE; 1012879888. MEDLINE; 1012879888888 MEDLINE; 1012879888 MEDLINE; 101287988 MEDLINE; 10128798 MEDLINE; 101	f; 637 AA. ced) sequence update) annotation update)

(MT)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Tabular output not generated. >US-09-331-631-24 (31-85) from US09331631.pep

Run on:

Sat May 13 08:59:23 2000; MasPar time 6.82 Seconds 380.618 Million cell updates/sec 12:0N (1) 235 T

Description: C Perfect Score: Sequence:

1 DEDDRRGGHSLQQCVQRCRQ.....DDQQQHGRHEQEEEQGRGRG 55

Post-processing: Minimum Match 0% Listing first 45 summaries

142080 segs, 47172406 residues

Searched:

Scoring table:

PAM 150 Gap 11

pir62
1:pir1 2:pir2 3:pir3 4:pir4

Database:

Statistics: Mean 32.380; Variance 69.376; scale 0.467

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Length	DB.	ID	Description	· z
ا بر	425	100.0	637	2	\$35221	lin Beg1	5e-6
N	174	40.9	122	2	C53234	lin-10 - maize	070
ω	174	40.9	582	N	B53234	n-like storage	<u>,</u> ,
4	170		236	N	T01662	,	٠,
ر د	168		407	N	T02258	inl - m	. 046
ייס	168		540	N	S21825	æ	۱,
7	168		573	N	5323	globulin-18, GLB18	۱,
3 0 ·	135		588		FWCNAB	lin B pre	1.000-00
۰ ع	131	30.8	509	N	S08059	in type	. 000
10	107		605	N	S06398	-	
1	105		957	N	D70835	glycin	0.256-04
12	101		1549	<u>, , , , , , , , , , , , , , , , , , , </u>	A40691	٠,	2.0/6-0
13	92	•	1021	N	T10748	mannan endo-1,4-Deta-	
14	92	•	1023	N	S12519	iru	
15	92	21.6	1898	—	A45973	man	4.40e-02
16	89	•	566	N	S22477	C	1 130-01
17	89	•	613	N	S27770	procein	1 130-01
cr cr	89		625	ν	A34615	n - rac	1 1 1
19	89		1200	N	T15921	prote	1 550.01
20	88		2248	۳	D42088	YCL	
21	87	20.5	330	N	S28102	capny	
22	85		450	N	G01158	osine ki	3.926.01
23	85	. 20.0	678	2	S54308	DNA binding protein -	0.926

44 5	41	39 40	37 38	35 6	မ ယ နှ	32 2	30	28	26	24
8383	8 CB W W	83 83	83	88	833	8 8	2000	000	200	0 00 - Ui
19.5 19.5 19.5			19.5 19.5	19.5 19.5	19.5	19.5	19.8	10.0		20.0
1218 1218 1920	837 1127	560	559	5 5 5	559	5 60 6	1255	524	380	788 325
NNN	2 -	22	22	202	N) N	ى ب ر) N I) N	N N
A2642/ A26588 A53188	A29512 A25018	S52943	S52940 S52950		S52954	S52941	A24571	JQ1730	\$36789 \$3640	I51530 JH0189
period clock protein pericentrin - mouse	circadian rhythm prot	period protein (clone	protein (protein (protein	protein			gene RD protein - hum class III histocompat	integrin beta-3 subun arginine/aspartate-ri
7.21e-01 7.21e-01	7.21e-01	7.21e-01	7.21e-01 7.21e-01	7.21e-01 7.21e-01	7.21e-01 7.21e-01	7.21e-01 7.21e-01	5.32e-01 7.21e-01	5.32e-01 5.32e-01	5.32e-01 5.32e-01	3.92e-01 5.32e-01

ACCESSIONS REFERENCE *authors *journal *title	RESULT 2 ENTRY TITLE ORGANISM	Db 31 DEDDRRC Qy 31 DEDDRRC	Query Match 100.0%; Best Local Similarity 100.0%; Matches 55; Conservative	FEATURE 174-190 SUMMARY	_position ICATION S	GENETICS	*accession S3521 **molecule_type mRNA **residues 1-63 **cross-references E	*cross-referen	a	ACCESSIONS REFERENCE	DATE	TITLE ORGANISM	RESULT 1 ENTRY
Hh.	C53234 #type complete globulin-10 - maize #formal_name Zea mays #common_name maize 02-May-1994 #sequence_revision 18-Nov-1994 #text_change	DEDDRRGGHSLQQCVQRCRQERPRYSHARCVQECRDDQQQHGRHEQEEEOGRGRG 85	100.0%; Score 425; DB 2; Length 637; larity 100.0%; Pred. No. 2.45e-60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	<pre>#product globulin Beg1 *status predicted *label MAT #length 637 *molecular-weight 72252 *checksum 3309</pre>	superfamily glycinin glycoprotein	Beq1	Coession S3921 ##molecule_type mRNA ##residues 1-637 ##Iabel HEC ##cross-references EMBL:M64372; NID:g167003; PID:g167004	constructed company mapping and regulation of expression. cross-references MUID:93287988	Gen. Genet. (1993) 239:20	N35221 S35221 Heck, G.R.; Chamberlain, A.K.; Ho, T.H.D.	03-Feb-1994 *sequence_revision U3-Feb-1994 *Lext_Change 26-Aug-1999	globulin Begi precursor - parrey #formal_name Hordeum vulgare #common_name barley	S35221 #type complete

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

HPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 08:40:44 2000: MasPar time 96.60 Seconds 18.915 Million cell updates/sec

Run on:

5501DM0:22

Tabular output not generated

Description: Perfect Score:

Sequence:

EDDNHHHHGGHKSGQCVRRC.....EKRQERSRHEADDRSGEGSS 60

Searched: Scoring table: PAM 150 Gap 11 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: swiss-prot38 1:swissprot

Statistics: Mean 34.513; Variance 68.998; scale 0.500

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

222 222 232 232 233	Result
88899999999999999999999999999999999999	Score
98.7 22.1 22.9 20.9 20.9 20.7 20.7 20.7 20.7 20.7 20.7 20.7 20.7	Query Match
573 605 1400 380 588 412 1200 1898 632 246 466 470 722 722 722 722 722 723 723 723 723 723	Length
	BB
GLB1_MAIZE VCLA_GOSHI TRHY_RABIT RDP_HUMAN VCLB_GOSHI YNP1_CAEEL DDX8_CAEEL TRHY_HUMAN VCRB_CAEEL CYP8_CAEEL C	ID
GLOBULIN-1 S ALLELE PR VICILIN GC72-A PRECURS TRICHOHYALIN RD PROTEIN PROTEIN VICILIN C72 PRECURSOR HYPOTHETICAL 49.5 KD P PUTATIVE PRE-MRNA SPLI TRICHOHYALIN DIHYDROPYRIDINE-SENSIT HYPOTHETICAL 28.5 KD P PEPPIDYL-PROLYL CIS-TR HYPOTHETICAL 28.5 KD P PEPPIDYL-ROLYL CIS-TR HYPOTHETICAL 28.5 KD P PEPPIDYL-ROLYL CIS-TR HYPOTEIN (WL633) 3-150PROPYLMALATE DEHY LEGUMIN A PRECURSOR LEGUMIN APROTEIN VND LEGUMIN APROTEN VND LEGUMIN APROTEN VND LEGUMIN APROTEN VND LEGUMIN APROTEN VND HYPOTHETICAL PROTEIN K VALTAGE-DEPENDENT N-TY THMEDIATE-EARLY PROTEIN K XANTHINE DEHYDROGENASE VOLTAGE-DEPENDENT N-TY VALTAGE-DEPENDENT N-TY THMEDIATE-EARLY PROTEIN K XANTHINE DEHYDROGENASE VOLTAGE-DEPENDENT N-TY	Description
1.62e-69 8.10e-04 8.10e-03 8.82e-03 1.23e-02 1.23e-02 1.23e-02 1.23e-02 3.33e-02 3.33e-02 4.63e-02 4.63e-02 6.41e-02 6.41e-02 6.41e-02 8.87e-02 8.87e-02	Pred. No.

EMBL; M24845; AAA33467.1; -. HSSP; P50477; 1CAW. MAIZEDB; 30181; -

; PF00546; Seedstore_7s; 1.
storage protein; Signal.

4	4	43	2	41	40	3 9	38	37	36 6	35 5	3 4	ω w	32	31	30	29	28	27	26	25	24	
ä	80	80	81	81	81	81	82	82	82	82	82	83	83	84	84	85	85	8 5	85	85	85	
1/.4	17.4	17.4	17.6	17.6	17.6	17.6	17.9	17.9	17.9	17.9	17.9	18.1	18.1	18.3	18.3	18.5	18.5	18.5	18.5	18.5	18.5	
TIO	612	587	1247	524	516	437	1227	1106	661	503	416	494	285	1121	382	935	815	668	344	197	196	
F	يسو ه	- ا	_	ب	_	ب	ب		μ	<u>س</u>	م سو	μ	۳,		<u>س</u>	<u>, , </u>	۳,	<u>س</u>		μ	۳	
OXA*_DOMAN	DYIZ_MOUSE	ESR1_POEGU	POLS_ONNVG	SBP_SOYBN	LEGB_GOSHI	YGY3_HALSQ	PR16_HUMAN	STC_DROME	F13B_HUMAN	YOW5_CAEEL	U2AF_DROME	SFR4_HUMAN	YD41_MYCLE	DDX8_ARATH	NCAP_CVCAE	RNE_HAEIN	MK07_HUMAN	F13B_MOUSE	SFR6_HUMAN	TRSF_DROME	SFR2_CAEEL	
	OFFICE INTERMEDIATE OF	ESTROGEN RECEPTOR (ER)	STRUCTURAL POLYPROTEIN	SUCROSE-BINDING PROTEI	LEGUMIN B PRECURSOR (B	HYPOTHETICAL 50.6 KD P	PRE-MRNA SPLICING FACT		Η	_	U	FACTOR, ARGIN	ICAL 30.9 KD P	-		E (EC 3.1	•••		CTOR, ARGIN	KANST		
4	1.500+00	1.506+00	1.10e+00	1.10e+00	1.10e+00	1.10e+00	8.12e-01	8.12e-01	8.12e-01	8.12e-01	8.12e-01	5.95e-01	5.95e-01	4.36e-01	4.36e-U1	3.18e-01	3.18e-01	3.18e-01	3.108-01	٠	٠	

3 6	38	ဂ္ဂ	റ്റ	င္ပ	ဂ	င္ပ	ဥ	ဌ	გ	ဂ္ဂ	ဂ ်	S	38	3 6	3 2	R	RT	æ	RX	Ŗ	R	RL	RŢ	X,	RA	RC	RΡ	RN	ဂ္ဂ	റ്റ	გ	လ	ဂ X	DE	ם ב	35	č	TD G	1
J	or send an email to license@isb-sib.cn).	ee http://w	modified and this statement is not removed. Usage by and for commercial	as its c	the European Bioinformatics Institute. There are no restrictions on its	between the Swiss Institute of Bioinformatics and the EMBL outstation -	This SWISS-PROT entry is copyright. It is produced through a collaboration		CONVICILIN, CONGLYCININ, ETC.).	-!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,		THE DESIGNATION L, I; AND S FOR LARGE, INTERMEDIATE, AND SMALL	-i- POLYMORPHISM: THE THREE MOST COMMONLY OCCURRING GLB1 ALLELES HAVE	STATE THREE TRANSPORTED TO THE ANGLATION DODING	BIOCHEM, GENEC. 4/14/14/2011/1911.	genes.";	"Characterization of embryo globulins encoded by the maize Glb		MEDLINE; 89374022.	SEQUENCE OF 87-100.	(2)	Plant Physiol. 91:636-643(1989).		"Molecular characterization of the major maize embryo globulin encoded		STRAIN-CV. INBRED LINE VA26;	SEQUENCE FROM N.A.	[1]	Poaceae; Zea.	euphyllophytes; Spermatophyta; Magnollophyta; Lillopsida; Poales;	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	Zea mays (Maize).		S ALLELE PI	(Rel.	л <u>г</u>		GLB1 MAIZE STANDARD; PRT; 573 AA.	

Statistics: Result Database: Post-processing: Searched: Scoring table: Sequence: Description: Perfect Scor Title: Tabular output MPsrch_pp ŏ 9 10 113 13 14 116 117 118 score and is Pred. No. Score Score: is Release 3.1A John F. Collins, Biocomputing Research Copyright (c) 1993-1998 University of Edinburgh, U Distribution rights by Oxford Molecular Ltd No. is the number of results predicted by chance to have a greater than or equal to the score of the result being printed, s derived by analysis of the total score distribution. protein · not generated Match PAM 150 Gap 11 sptrembil2
1:sp_AarChea 2:sp_bacteria 3:sp_fungi 4:sp_human
1:sp_arChea 2:sp_bacteria 7:sp_mhc 8:sp_organelle
5:sp_inage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus Mean 33.511; Variance 71.985; Minimum SUS-09-331-631-22 (25-84) from US09331631 Sat May 13 08:42:53 2000; Listing 225878 seqs, EDDNHHHHGGHKSGQCVRRC.....EKRQERSRHEADDRSGEGSS Length protein database search, using Smith-Waterman algorithm Match 0% first 45 summaries В 69334122 residues 003865
0013856
0013856
0013856
0013856
0013856
0013856
001249
081257
081257
081256
081256
081259
081256 H SUMMARIES . pep GLOBULIN-1 (FRAGMENT).
EMBRYO GLOBULIN.
HYPOTHETICAL 45.3 KD P
EARLY EMBRYOGENESIS PR GLOBULIN-1
GLOBULIN-1
GLOBULIN-1 MasPar time 167.03 Seconds 24.906 Million cell updates/sec Description scale 0.466 Research Unit. burgh, U.K. 1 0 ALLELE PR
1 L ALLELE PR
1 (FRAGMENT)
1 S ALLELE PR
1 (FRAGMENT)
1 (FRAGMENT) (FRAGMENT).
(FRAGMENT).
(FRAGMENT).
(FRAGMENT). (FRAGMENT). (FRAGMENT). (FRAGMENT). (FRAGMENT). (FRAGMENT). Pred. 2.54e-63 2.54e-63 2.54e-63 9.31e-63
2.21e-63
2.21e-63
3.42e-59
4.64e-50
2.15e-48
1.05e-44 4.50e-64 3.03e-15 5.82e-05 1.01e-03 .08e-66 ,92e-63 .22e-64 .22e-64 60 25012405.

(4			
4 4 4 4 10 W 4 70	440 120 120 120 120 120 120 120 120 120 12	3 3 3 3 3 3 3 4 3 5 4 3 5 4 3 5 4 3 5 4 3 5 4 3 5 4 3 5 4 3 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5	32 32 32	22 22 24 26
88 88 87	88888 9999	000000	91 91 91 91	103 99 97 97 96
19.2			19.8 19.8 19.8	
515 573 606			292 10 292 10 342 11 450 4 499 10	
5 045896 10 Q9ZR40 4 Q9Y341				
ROTEIN. JXILIARY F ANNEL BETA	COSMID K09H11. INITIATION FACTOR IF2. CODED FOR BY C. ELEGAN SIMILAR TO AXONEME-ASS	LET 858. SERINE-THREONINE KINAS NUCLEIC ACID BINDING P PUTATIVE COINTEGRATE R 45KDA SPLICING FACTOR.	CLEAR RIBON S III REGIO KINASE ACT (FRAGMENT).	THROMBOSPONDIN-RELATED SBCFK-1 PROTEIN PRECUK F55A11.6 PROTEIN. F28C1.1 PROTEIN. PLASMID PPL1, COMPLETE FCGI-74 PROTEIN. FOR SUBBLE AUTHITABY FAC
440	3.25e-01 N 3.25e-01 S 3.25e-01 S 3.25e-01	3.25e	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	3.79e-03 1.39e-03 2.65e-02 2.65e-02 2.65e-02 6.86e-02

```
RESULT TO SOLUTION OF THE PROPERTY OF THE PROP
RESULT
ID QC
AC QC
DT 01
                                                                                                                                                                                  ş
                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 6
                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q03863 PRELIMINARY; PRT: 122 AA. Q03863; 01-NOV-1996 (TrEMBLrel. 01, Created) 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) 01-NOV-1999 (TrEMBLrel. 12, Last annotation update) GLOBULIN-1 0 ALLELE PRECURSOR (GLB1-0) (7S-LIKE).
IT 2
003865 PRELIMINARY
003865;
01-NOV-1996 (TremBirel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X59085; CAA41811.1; -.
MENDEL; 14637; Zeama;1188;14637.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genetics 129:863-872(1991).
-T-MISCELLANEOUS: THREE PROTEIN-PROCESSING STEPS OCCUR IN THE FORMATION OF THE MATURE PROTEIN FROM THE PRIMARY TRANSLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE; 92090707.
BELANGER F.C., KRIZ A.L.;
"Molecular basis for alle.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zea mays (Maize).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnollophyta; Liliopsida; Poales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLB1-0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seed storage protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Poaceae;
                                                                                                                                                                               25 EDDNHHUHGGHKSGQCVRRCEDRPWHQRPRCLEQCREEEREKRQERSRHEADDRSGEGSS 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, ETC.).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRODUCT
                                                                                                                                                                                                                                                                                                60;
                                                                                                                                                                                                                                                                                                                                            n 100.0%;
Simblarity 100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                 20 ]
122 AA;
                                                                                                                             ٠,
                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               for allelic polymorphism of the maize Globulin-1
                                                                                                                                                                                                                                                                                                                                                                                                                                       122 (
13902 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19
      01,
                                                                                                                                                                                                                                                                                                                                            Score 459; DB 10;
Pred. No. 9.08e-66;
         Created
                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
GLOBULIN-0.
; 89C81691 CRC32;
                                                                                                                                                                                                                                                                                                                0
                                                                        PRT;
                                                                                                                                                                                                                                                                                                                      Miamatches
                                                                        582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ₹
                                                                        ζ
                                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                                                                           Length 122;
                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                         0
```

****** (ME)

Release 3.1A John F. Collins, Biocomputing Research Unit. Copyright (c) 1993-1998 University of Edinburgh, U.K. Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Sat May 13 08:46:12 2000; MasPar time 6.97 Seconds 406.080 Million cell updates/sec

-25010NO.22

Description: C >US-09-331-631-22 (25-84) from USD9-31631.pep

Tabular output not generated

Run on:

EDDNHHHHGGHKSGQCVRRC....EKRQERSRHEADDRSGEGSS 60

Scoring table: PAM 150 Gap 11 Sequence:

Searched: 142080 segs, 47172406 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database:

pir62
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 33.532; Variance 75.335; scale 0.445

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description	e d
N 14	459 459	100.0	122	NN	C53234 B53234	globulin-10 - maize vicilin-like storage	5.93e-ti
w	453		540	N	S21825	like storag	.14e-
•	453	98.7	573	N	w	-1S,	e-6
5	446	7	407	N	25	obulin1 - m	ø
6	343	74.7	236	N	T01662	globulin-1 - maize (f	2.79e-42
7	180	39.2	637	N	S35221	globulin Begl precurs	1.25e-14
80	106		605	N	S06398	bulin	3e-
9	103		1407		S28589	,	e-0
10	96		325	N	JH0189	arginine/aspartate-ri	e
11	96	•	380	2	S36789	gene RD protein - hum	.12e-
12	96	20.9	382	N	A33640	ii s	ė
13	96	•	588	_	FWCNAB	alpha-globulin B prec	.12e-
14	95		346	N	S41001	2	5.58e-02
15	95	20.7	1200	2	T15921		. 58e
16	95	20.7	1898	<u>, , , , , , , , , , , , , , , , , , , </u>	A45973	trichohyalin - human	5.58e-02
17	94		606	N	S21046	calcium channel prote	.55e
18	94	20.5	632	N	S21048		7.55e-02
19	92	•	241	N	S44893	ZK1236.7 protein - Ca	1.38e-01
20	92	20.0	466	ν	T15877	hypothetical protein	. 38e
21	91	9	33	N	A41822	\vdash	1.85e-01
22	16	•	375	N	A40112	MHC-region RD-repeat	1.85e-01
23	91	19.8	450	N	601158	tyrosine kinase activ	1.85e-01

4 5	44	43	42	41	40	39	38	37	36			ω ω	32	u L	30	29	28	27	26	25	24
8 5	85	85	85	85	85	85	85	86	86	86	86	87	89	89	99	89	90	90	90	90	90
18.5	18.5	18.5	18.5	18.5	18.5	18.5	18.5	18.7	18.7	18.7	18.7	19.0	19.4	19.4	19.4	19.4	19.6	19.6	19.6		19.6
2339	951	815	668	566	344	197	196	2336	1473	1280	509	1335	1276	471	407	163	722	520	517	507	203
N	ب	N	N	N	2	N	N	2	N	2	N	N	N	N	μ.	N	N	N		N	N
A42566	E64066	B56708	A46013	S22477	S59043	A29648	T15917	A45386	T13855	T00365	S08059	S07245	T02711	S02016	EDBEQ3	I38191	S57246	S08237	FWPMLA	T06452	S09559
omega-conotoxin-sensi	ribonuclease E (EC 3.	extracellular signal-	coagulation factor XI	vicilin precursor - c	spliling factor SRp55	female-specific trans	hypothetical protein	omega-conotoxin-sensi	suppressor of sable p	hypothetical protein	alpha-globulin type B	xanthine dehydrogenas	hypothetical protein	Ul snRNP 70K protein	immediate-early prote	nucleic acid binding	ventral nervous syste	2	legumin A precursor	probable legumín A pr	legumin A (clone pCD4
1.07e+00	1.07e+00	8.03e-01			8.03e-01					3.35e-01		•	•	2.49e-01	2.49e-01						

ALIGNMENTS

RESULT 2 B53234 *type complete ENTRY vicilin-like storage protein Glb1-L, embryo - maize ALTERNATE_NAMES: globulin-lL ORGANISM *formal_name Zea mays *common_name maize O2-May-1994 *sequence_revision 18-Nov-1994 *text_change 02-May-1997 ACCESSIONS B53234; S21824 REFERENCE #authors Belanger, F.C.; Kriz, A.L. *bournal Genetics (1991) 129:863-872 *title Molecular basis for allelic polymorphism of the maize @cross-references MUID:92090707 @accession B53234 B53234	Query Match 100.0%; Score 459; DB 2; Length 122; Best Local Similarity 100.0%; Pred. No. 5.93e-63; Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Db 25 EDDNHHHIGGHKSGOCVRRCEDRPWHQRPRCLEQCREEEREKRQERSRHEADDRSGEGSS 84	#cross-references MUID-92090707 #contents Black Beauty popcorn #accession C5324 ##status preliminary ##nolecule_type DNA ##cross-references EMBL:X59085; NID:g22281; PID:g22282 ##note sequence extracted from NCBI backbone (NCBIN:71282, NCBIP:71286) SUMMARY #length 122 #molecular-weight 13902 #checksum 4000	TITLE globulin-10 - maize ORGANISM *tope complete *formal_name Zea mays *common_name maize DATE 02-May-1994 *sequence_revision 18-Nov-1994 *text_change 10-Oct-1997 ACCESSIONS C53234; S21823 REFERENCE A53234 **journal #title Genetics (1991) 129:863-872 # delanger, F.C.; Kriz, A.L. # delanger, F.G.; Kriz,
--	--	--	---

tage t